



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 113398

TO: Michael Borin
Location: rem/2a55/2c70
Art Unit: 1631
Wednesday, February 04, 2004

Case Serial Number: 09/582486

From: Toby Port
Location: Biotech-Chem Library
Remsen 1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Borin,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:33:58 ; Search time 21 Seconds

(without alignments)

1424.212 Million cell updates/sec

Title: US-09-582-486-1

Perfect score: 1637

Sequence: 1 MDTTPTFSLAELOOGLHOD.....ATFQDWIGNYNIRRTSKA 311

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96158682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1629	99.5	311	2 T52312	deacetoxycephalosporin
2	1627	99.4	311	2 A32043	deacetoxycephalosporin
3	1185.5	72.5	314	2 S30900	deacetoxycephalosporin
4	965.5	59.0	310	2 S40253	deacetoxycephalosporin
5	939.5	57.4	318	2 A39204	deacetoxycephalosporin
6	937	57.2	319	2 S54100	deacetoxycephalosporin
7	926.5	56.6	332	2 A29711	deacetoxycephalosporin
8	852	52.0	313	2 S54101	deacetoxycephalosporin
9	176	10.8	329	2 A61155	isopenicillin N sy
10	164	10.0	329	2 A29894	isopenicillin N sy
11	162	9.9	326	2 S08218	isopenicillin N sy
12	162	9.9	326	2 S54099	isopenicillin N sy
13	146.5	8.9	362	2 D84713	probable dioxigena
14	144.5	8.8	376	2 D96635	probable gibberell
15	143	8.7	333	2 B32344	isopenicillin N sy
16	142.5	8.7	358	2 C84713	probable dioxigena
17	140	8.6	348	2 S33510	flavonol synthase
18	138	8.4	328	2 S15284	isopenicillin N sy
19	137	8.4	321	2 A58458	isopenicillin N sy
20	136.5	8.3	380	2 T48491	gibberellin 20-oxi
21	134.5	8.2	349	2 T07373	flavonol synthase
22	133.5	8.2	331	2 A26467	isopenicillin N sy
23	131.5	8.0	355	2 AG0160	probable iron/asco
24	129.5	7.9	331	2 S04441	isopenicillin N sy
25	125	7.6	382	2 T11847	gibberellin 20-oxi
26	124.5	7.6	370	2 T11849	gibberellin 20-oxi
27	123	7.5	339	2 T05903	iron deficiency pr
28	122.5	7.5	331	2 A27355	isopenicillin N sy
29	122.5	7.5	386	2 T09664	gibberellin 20-oxi

30	122.5	7.5	386	2 T09675	probable gibberell
31	121	7.4	356	2 T05119	leucoanthocyanidin
32	121	7.4	371	2 T08008	leucoanthocyanidin
33	118.5	7.2	377	2 T10222	gibberellin 20-oxi
34	118	7.2	365	2 S31921	naringenin 3-dioxy
35	117.5	7.2	358	2 S44261	SRG1 protein - Ara
36	117.5	7.2	370	2 T47932	1-aminocyclopropan
37	117	7.1	320	2 C83628	probable oxidoredu
38	116.5	7.1	356	2 T05551	SRG1 protein-relat
39	115	7.0	380	2 T06787	gibberellin 20-oxi
40	114.5	7.0	321	2 T41002	hypothetical prote
41	114	7.0	356	2 T07972	leucoanthocyanidin
42	113.5	6.9	338	2 S47972	dioxygenase-iron
43	113	6.9	298	2 T09145	ethylene-forming e
44	113	6.9	352	2 T01606	probable flavonol
45	113	6.9	380	2 T06439	gibberellin 20-dio

ALIGNMENTS

RESULT 1

T52312

deacetoxycephalosporin C synthetase [imported] - Streptomyces clavuligerus

C:Species: Streptomyces clavuligerus

C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 24-Aug-2001

C:Accession: T52312

R:Kovacevic, S.; Tobin, M.B.; Miller, J.R.

J. Bacteriol. 172, 3952-3958, 1990

A:Title: The beta-lactam biosynthetic genes for isopenicillin N epimerase and deacetoxyce

A:Reference number: Z26033; MUID:90299822; PMID:1694525

A:Accession: T52312

A:Status: Preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-311 <KOV>

A:Cross-references: EMBL:M32324; PIDN:AAA26715.1

C:Genetics:

A:Gene: cefE

C:Superfamily: isopenicillin N synthase

Query Match 99.5%; Score 1629; DB 2; Length 311;
Best Local Similarity 99.7%; Pred. No. 4.9e-140;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MDTTPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVIDFFEHGSE	60
Db	1	MDTTPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVIDFFEHGSE	60
Qy	61	AEKRAVTSVPVTTMRGFTGLESESTAQITNTGSYSVSMCYSMTADNLPSPGDFRIWT	120
Db	61	AEKRAVTSVPVTTMRGFTGLESESTAQITNTGSYSVSMCYSMTADNLPSPGDFRIWT	120
Qy	121	QYFDQVYTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM	180
Db	121	QYFDQVYTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM	180
Qy	181	APHYDLSWMTLIOOTPCANGFVSLQAEVGAFTDLPIYRDAVLVFCGATATLVGTQGVKA	240
Db	181	APHYDLSWMTLIOOTPCANGFVSLQAEVGAFTDLPIYRDAVLVFCGATATLVGTQGVKA	240
Qy	241	PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGG	300
Db	241	PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGG	300
Qy	301	NYVNIRRTSKA 311	
Db	301	NYVNIRRTSKA 311	

RESULT 2

A32043

deacetoxycephalosporin C synthetase - Streptomyces clavuligerus

C:Species: Streptomyces clavuligerus

THIS PAGE BLANK (USPTO)

C;Date: 13-Jul-1989 #sequence_revision 13-Jul-1989 #text_change 19-May-2000
C;Accession: A32043
R;Kovacevic, S.; Weigel, B.J.; Tobin, M.B.; Ingolia, T.D.; Miller, J.R.
J. Bacteriol. 171, 754-760, 1989
A;Title: Cloning, characterization, and expression in *Escherichia coli* of the Streptomycin
A;Reference number: A32043; MUID:89123150; PMID:2644235
A;Accession: A32043
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <COV>
C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 99.4%; Score 1627; DB 2; Length 311;
Best Local Similarity 99.4%; Pred. No. 7.5e-140;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVTPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFEHGSE 60
DB 1 MDVTPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFEHGSE 60

QY 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGSDYSYSCMYSGMTADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGSDYSYSCMYSGMTADNLPSPGDFGRITW 120

QY 121 QYFDRQYTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPVPOVPEHRSAAEQPLRM 180
DB 121 QYFDRQYTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPVPOVPEHRSAAEQPLRM 180

QY 181 APHYDLSNVTLTIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSNVTLTIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKA 240

QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300

QY 301 NYVNIIRTSKA 311
DB 301 NYVNIIRTSKA 311

RESULT 3
S30900
deactoxycephalosporin C synthetase - Streptomycetes lactamdurans
C;Species: Streptomycetes lactamdurans
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-2000
C;Accession: S30900; S28383
R;Coque, J.J.R.; Martin, J.F.; Liras, P.
Mol. Gen. Genet. 236, 453-458, 1993
A;Title: Characterization and expression in Streptomycetes lividans of cefD and cefE genes
clavuligerus.
A;Reference number: S30900; MUID:93173127; PMID:8437592
A;Accession: S30900
A;Molecule type: DNA
A;Residues: 1-314 <COV>
A;Cross-references: EMBL:213974; NID:949299; PIDN:CAA78376.1; PID:949300
A;Note: the source is designated as Nocardia lactamdurans
C;Genetics:
A;Gene: cefE
C;Superfamily: isopenicillin N synthase

Query Match 72.5%; Score 1186.5; DB 2; Length 314;
Best Local Similarity 69.8%; Pred. No. 7e-100;
Matches 217; Conservative 41; Mismatches 52; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFEHGSE 60
DB 3 DATVPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFEHGTE 62

QY 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGSDYSYSCMYSGMTADNLPSPGDFGRITW 120
DB 63 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGSDYSYSCMYSGMTADNLPSPGDFGRITW 122

QY 121 QYFDRQYTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPVPOVPEHRSAAEQPLRM 180
DB 123 DYFARMYASQDVARQVLTSSVGAEPVQMDAFDCEPLLRFRYPVPOVPEHRSAAEQPLRM 182

QY 181 APHYDLSNVTLTIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKA 240
DB 183 APHYDLSNVTLTIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKA 242

QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300
DB 243 PGHVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 302

QY 301 NYVNIIRTSKA 311
DB 303 NYINIRKATAA 313

RESULT 4
S40253
deactoxycephalosporin C synthetase - Streptomycetes lactamdurans
C;Species: Streptomycetes lactamdurans
C;Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 19-May-2000
C;Accession: S40253
R;Coque, J.
submitted to the EMBL Data Library, February 1993
A;Reference number: S40253
A;Accession: S40253
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-310 <COV>
A;Cross-references: EMBL:221687; NID:9438193; PIDN:CAA79803.1; PID:9438194
A;Note: the source is designated as Nocardia lactamdurans
C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 59.0%; Score 965.5; DB 2; Length 310;
Best Local Similarity 61.2%; Pred. No. 7.8e-80;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFEHGSEA 61
DB 3 DKTVPVFSMAELDGSQRQDEFREWAR-RGVFYLTYGATERDHRVATDTAMDFFAQGTAE 61

QY 62 EKRAVTSVPVPTMRGFTGLESESTAQITNTGSDYSYSCMYSGMTADNLPSPGDFGRITW 121
DB 62 EKQAVTIVPTMRGYSALAEASTAQVNTGTVDYSMSYSGIGGLNLPSPKEFESVWTD 121

QY 122 YFDRQYTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPVPOVPEHRSAAEQPLRMA 181
DB 122 YFDSLRYAAQETARLVLTAGTYDGDLDLTLLDCDPLVLRYPVPEHRAAEYEPERMA 181

QY 182 PHYDLSNVTLTIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKA 241
DB 182 PHYDLSNVTLTIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKA 241

QY 242 RHVVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 301
DB 242 NHHVSPDAMLGSDRTSSVFFLRPNADFTFVSPDARKYGLDVSMDMEKATFGDWIGTN 301

QY 302 YVNI 305
DB 302 YVTM 305

RESULT 5
A39204
deactoxycephalosporin C synthetase hydroxylase - Streptomycetes clavuligerus
C;Species: Streptomycetes clavuligerus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-May-2000
C;Accession: A39204; A23713
R;Kovacevic, S.; Miller, J.R.
J. Bacteriol. 173, 398-400, 1991
A;Title: Cloning and sequencing of the beta-lactam hydroxylase gene (ceff) from Streptomycetes.

THIS PAGE BLANK (USPTO)

Reference number: A39204; MUID:91100311; PMID:1987130

Accession: A39204

Molecule type: DNA

Residues: 1-318 <KOV>

Cross-references: GB:M3809; GB:M37186; NID:g153206; PIDN:AAA26716.1; PID:g153207

Baker, B.J.; Dotzlaef, J.E.; Yeh, W.K.

Biol. Chem. 266, 5087-5093, 1991

Title: Deacetoxycephalosporin C hydroxylase of Streptomyces clavuligerus. Purification

Reference number: A23713; MUID:91161600; PMID:2002049

Accession: A23713

Molecule type: protein

Residues: 2-29,92-100 <BAK>

Genetics:

Gene: cefF

Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 57.4%; Score 939.5; DB 2; Length 318;
Best Local Similarity 58.6%; Pred. No. 1.8e-77;
Matches 180; Conservative 43; Mismatches 81; Indels 3; Gaps 1;

2 DTTVPFSLAELOQGLHODEFRRLRDKGLFYLTDCGLTDTLKSADLVDPFHEGSEA 61

3 DTPVPFENLAUREGADQKRECVTGMGVFYLTVGAGDKDRLATDTAMDFFANGTEA 62

62 EKRAVTSVPVPMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLPFGSGDFGRITWQ 121

63 EKAAVTDVPMRGVSALAEASTAQVTRTGYTSDYSGMTADNLPFGSGDFGRITWQ 122

122 YFDQVYTSRAVAREVLRTG---TEPDGGVAFDCEPLLRFPPQVPEHRSAAEQPL 178

123 YFDKLYAAQETARLVLTASGGYDAEIVGSLDELDDADPVLRLRYPEVPEHRSAAHPR 182

179 RNAPHYDLSMTVLIQOTPCANGFVSLOAEVGAFTDLPYRPDAVLVFCGAIATLVGGQV 238

183 RNAPHYDLSMTVLIQOTPCANGFVSLOAEVGAFTDLPYRPDAVLVFCGAIATLVGGQV 242

239 KAPRHVAAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWI 298

243 PAPRHVRSFGAGMRGSDRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWI 302

299 GGNVNI 305

303 GTNYVTM 309

RESULT 6

leacetoxycephalosporin C synthetase 1 - Lysobacter lactangenus (strain YK90)

Species: Lysobacter lactangenus

Variety: strain YK90

Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-May-2000

Accession: S54100

Kimura, H.; Iwata, M.; Miyashita, H.; Shimizu, Y.; Sumino, Y.; Suzuki, M.

submitted to the EMBL Data Library, October 1990

Description: Gene cluster involved in the cephalosporin biosynthesis from Lysobacter

Reference number: S54099

Accession: S54100

Molecule type: DNA

Residues: 1-319 <KIM>

Cross-references: EMBL:X56660; NID:g769806; PIDN:CAA39984.1; PID:g769808

Experimental source: strain YK90

Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Keywords: antibiotic biosynthesis

Query Match 57.2%; Score 937; DB 2; Length 319;

Best Local Similarity 58.0%; Pred. No. 3.1e-77;

Matches 177; Conservative 45; Mismatches 83; Indels 0; Gaps 0;

2 DTTVPFSLAELOQGLHODEFRRLRDKGLFYLTDCGLTDTLKSADLVDPFHEGSEA 61

3 DSGIQIFDLDEHGVRLDSFRKLFERGVYFVREDDSIKTEHAKAMDAVMDLFGNSAE 62

62 EKRAVTSVPVPMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLPFGSGDFGRITWQ 121

63 QKALRNLTNVRGFSDLERASTARKGGETYDSWYISGLTDLNLFSPAFALWTG 122

122 YFDQVYTSRAVAREVLRTGTEPDGGVAFDCEPLLRFPPQVPEHRSAAEQPLRMA 181

123 YFDQVYTSRAVAREVLRTGTEPDGGVAFDCEPLLRFPPQVPEHRSAAEQPLRMA 182

182 PHYDLSMTVLIQOTPCANGFVSLOAEVGAFTDLPYRPDAVLVFCGAIATLVGGQVAP 241

183 PHYDLSMTVLIQOTPCANGFVSLOAEVGAFTDLPYRPDAVLVFCGAIATLVGGQVAP 242

242 RHVAAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGN 301

243 RHVAAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGN 302

302 YVNIR 306

303 YVNIR 307

RESULT 7

A29711

deacetoxycephalosporin C synthetase - fungus (Acremonium sp.)

Alternates names: expandase; hydroxylase

Species: Acremonium sp.

Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 19-May-2000

Accession: A29711; A41864

Samson, S.M.; Dotzlaef, J.E.; Slisz, M.L.; Becker, G.W.; Van Frank, R.M.; Veal, L.E.; Ye

Bio/Technology 5, 1207-1214, 1987

Title: Cloning and expression of the fungal expandase/hydroxylase gene involved in cepi

Reference number: A29711

Accession: A29711

Status: not compared with conceptual translation

Molecule type: DNA

Residues: 1-332 <SAM>

Note: the source is designated as Cephalosporium acremonium

Gutierrez, S.; Velasco, J.; Fernandez, F.J.; Martin, J.F.

J. Bacteriol. 174, 3056-3064, 1992

Title: The cefG gene of Cephalosporium acremonium is linked to the cefEF gene and encod

Reference number: A41864; MUID:92234966; PMID:1569032

Accession: A41864

Status: preliminary

Molecule type: DNA

Residues: 1-76, 'L' <GUT>

Experimental source: strain C10

Note: sequence extracted from NCBI backbone (NCBI:104773, NCBI:97574); this ORF is nc

Note: the source is designated as Cephalosporium acremonium

Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 56.6%; Score 926.5; DB 2; Length 332;

Best Local Similarity 56.7%; Pred. No. 2.9e-76;

Matches 177; Conservative 43; Mismatches 91; Indels 1; Gaps 1;

1 MDTTVPFSLAELOQGLHODEFRRLRDKGLFYLTDCGLTDTLKSADLVDPFHEGSEA 60

1 MTSKVPFRLDGLSKGKVLTELAEAVTKGIFLYTESGLVDDHTSARETCVDFPKNGSE 60

61 EKRAVTSVPVPMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLPFGSGDFGRITW 120

61 EKRAVTLADRNRGRFSALEWESTAVVTETGKYSYDSTCYSMGIGNLFNRRGDFVWQ 120

121 QYFDQVYTSRAVAREVLRTGTEPDG-GVEAFDCEPLLRFPPQVPEHRSAAEQPLR 179

121 DYFDQVYTSRAVAREVLRTGTEPDG-GVEAFDCEPLLRFPPQVPEHRSAAEQPLR 180

180 MAPHYDLSMTVLIQOTPCANGFVSLOAEVGAFTDLPYRPDAVLVFCGAIATLVGGQVAP 239

181 MAPHYDLSMTVLIQOTPCANGFVSLOAEVGAFTDLPYRPDAVLVFCGAIATLVGGQVAP 240

240 APRHVAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIG 299

241 APRHVAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIG 300

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)


```
QY 300 GNYVNIIRRTSKA 311
|||||
Db 301 GNYVNIIRDRKPA 312

RESULT 8
S54101
deacetoxycephalosporin C synthetase 2 - Lysobacter lactamgenus (strain YK90)
C:Species: Lysobacter lactamgenus
A:Variety: strain YK90
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-2000
R:Kimura, H.; Izawa, M.; Miyashita, H.; Shimizu, Y.; Sumino, Y.; Suzuki, M.
submitted to the EMBL Data Library, October 1990
A:Description: Gene cluster involved in the cephalosporin biosynthesis from Lysobacter
A:Reference number: S54099
A:Accession: S54101
A:Molecule type: DNA
A:Residues: 1-313 <KIM>
A:Cross-references: EMBL:X56660; NID:g769806; PIDN:CAA39985.1; PID:g769809
R:Experimental source: strain YK90
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase
C:Keywords: antibiotic biosynthesis

Query Match 52.0%; Score 852; DB 2; Length 313;
Best Local Similarity 52.9%; Pred. No. 1.6e-69;
Matches 165; Conservative 48; Mismatches 91; Indels 8; Gaps 2;

QY 2 DTTVPTFSLAEOGLHODFRCLRDKGLFYLTDCGLTDTLTKLSAKD---LVIDFFEH 57
|||||
Db 3 DRTVPTFALEQLQGEHKAALRCAMEKGAFLQ---IDGPLSAAHADRAGKAAMDFAT 58
|||||

QY 58 GSEAEKRAVTSVPVPTMRGFTGLESESTAQTNTGYSYSDYSCYSGMTADNLPFSGDQGR 117
|||||
Db 59 ADDAAQAVINANPTNRGFSPLGSESTACTNTGYSYVAMYSIGISNIPPTAHPR 118
|||||

QY 118 IMTOYFDROYTASRAVAREVLRTATGTPDGGVEAFDCEPLLRFRYPQVPEHRSABEQ 177
|||||
Db 119 LMSDFDLYYIGISQAAARAVLESMDVHLNTDIDALDCDPLVLRYPFDPVEDRCAEQP 178
|||||

QY 178 LRNAPHYDLSMVTLIQOQTPCANGFVSLOAEVGAFTDLPYRPDAVLVFCGATATLTGQ 237
|||||
Db 179 NRNAPHYDLSIVSLIQLQPCNGFVSLQVEIDGRFVEVPRPCGVVFCGSIAPLVSDGK 238
|||||

QY 238 VKAPRHVAAAPRDQIAGSRTSSVFLRNADFTFVPLARCGFDVSLDGETATFODW 297
|||||
Db 239 IKAPQHVVSFGAAQIGSNRTSSVFLRFXSEFSVPLAKALGMDGLDVGERTATFGW 298
|||||

QY 298 ICGYVNIIRTS 309
|||||
Db 299 CGANYMEMHVT 310
|||||

RESULT 9
A61155
isopenicillin N synthase (EC 1.14.11.-) pcbC [similarity] - Streptomyces griseus (strain
C:Species: Streptomyces griseus
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 26-May-2000
A:Accession: A61155
R:Garcia-Dominguez, M.; Liras, P.; Martin, J.F.
Antimicrob. Agents Chemother. 35, 44-52, 1991
A:Title: Cloning and characterization of the isopenicillin N synthase gene of Streptomyces
avuligerus.
A:Reference number: A61155; MUID:91197089; PMID:1901702
A:Accession: A61155
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-329 <GAR>
A:Cross-references: GB:X54609; NID:g509097; PIDN:CAA38431.1; PID:g509098
C:Genetics:
A:Gene: pcbC
C:Superfamily: isopenicillin N synthase
C:Keywords: antibiotic biosynthesis; ascorbic acid; iron; metalloprotein; oxidoreductase
```

```
F:48,212,268/Binding site: iron (His) #status predicted

Query Match 10.8%; Score 176; DB 2; Length 329;
Best Local Similarity 24.5%; Pred. No. 3.6e-08;
Matches 73; Conservative 39; Mismatches 126; Indels 60; Gaps 12;

QY 5 VPTFSLAELQOGLHOD-----EPRCLRDKGLFYLTDCGLTDTLTKLSAKDLVIDFFEHG 58
|||||
Db 10 VPTIDISPLSGDADDDKRVQAEINKACRESGFFVASHHGI---DVQLLKDVVNEFHRM 66
|||||

QY 59 SEAEK-----RAVTSVPVPTMRGFTGLESESTAQTNTGYSYSDYSCYSGMTADNLPF-- 111
|||||
Db 67 TDEEYDLAINAYNNKNNPRTRNGY-----YMAVKGKAVESWCYLNPSFSBDHPQI 117
|||||

QY 112 -SG---DFGRW-----TOYFDROYTASRAVAREVLRTATGTPDGGVEA--- 151
|||||
Db 118 RSGTPMHEGNTWPEKRRHRRFPCEQYRYDVSLSKVLMRGFALALG-KPEDFFDASLS 176
|||||

QY 152 ---FLDCEPLLRFRYPQVPEHRSABE-QPLRMAPHYDLSMVTLIQOQTPCANGFVSLOAE 207
|||||
Db 177 LADTILSAVTLIHYPVLEDPYPPVKTCPGDGTCLSFEDHLDVSMITVLFTQEVQN---LQVE 232
|||||

QY 208 VGGAPTDLPRPDVAVLFCGAIATLVTGGQVAPRRHVAAAPRRDQIAGSRTSSVFFL 265
|||||
Db 233 TADGQDLPTSGENFLVNCGTMYGTYLNDYFPAPNHRV-----KFINAERLSLPFFL 284
|||||

RESULT 10
A29894
isopenicillin N synthase (EC 1.14.11.-) [similarity] - Streptomyces clavuligerus
C:Species: Streptomyces clavuligerus
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 26-May-2000
A:Accession: A29894
R:Leskiw, B.K.; Aharonowitz, Y.; Mevarech, M.; Wolfe, S.; Vining, L.C.; Westlake, D.W.
Gene 62, 187-196, 1988
A:Title: Cloning and nucleotide sequence determination of the isopenicillin N synthet
A:Reference number: A29894; MUID:88212175; PMID:3130293
A:Accession: A29894
A:Molecule type: DNA
A:Residues: 1-329 <LES>
A:Cross-references: GB:M19421; NID:g153328; PIDN:AAA26770.1; PID:g153329
C:Superfamily: isopenicillin N synthase
C:Keywords: antibiotic biosynthesis; ascorbic acid; iron; metalloprotein; oxidoreduct
F:48,212,268/Binding site: iron (His) #status predicted

Query Match 10.0%; Score 164; DB 2; Length 329;
Best Local Similarity 23.8%; Pred. No. 4.4e-07;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

QY 5 VPTFSLAEL-----QOGLHQDEPRCLRDKGLFYLTDCGLTDTLTKLSAKDLVIDFFEHG 58
|||||
Db 10 VPTIDISPLFCGTDAARKVAEETHGACRSGSFFVATNHGV---DVQQLQDVVNEFHGM 66
|||||

QY 59 SEAEK-----RAVTSVPVPTMRGF-----TGLESSTAQTNTGYSYSDYSCYSGMT---A 106
|||||
Db 67 TDQEKHDLIAHAYNPDPNPHVNGYKAVPGRKAVESFCYLNPDCEGHPM-IAAGTPMHE 125
|||||

QY 107 DNLFPSCD-----FGRITQYFDROYTASRAVAREVLRTATGTPDGGVEA-----FLDCE 156
|||||
Db 126 VNLMPDEERHPRFRPFCEGYRYQMLKUSTVLMRGLALALG-RPEHFFDAALAEQDLSLV 184
|||||

QY 157 PLLRFRYPQVPEHRSABE-QPLRMAPHYDLSMVTLIQOQTPCANGFVSLOAEVGAFTDL 215
|||||
Db 185 SLIRYPVLEEVPPVKTGPDGQLLSFEDHLDVSMITVLFTQVQN-----LOVETVDGWRDI 240
|||||

QY 216 PYRPDAVLVFCGAIATLVTGGQVAPRRHVAAAPRRDQIAGSRTSSVFFLRPNADFTFSV 275
|||||
Db 241 PTSENDLFLVNCGTMYMAHVTNDYFPAPNHRV-----KFINAERLSLPFFLNGCGEAVIE- 293
|||||

QY 276 PLARECGFDVSLDGETATFQDWI 298
|||||
Db 294 PFVPE-GASEEVRNEALSVDYL 315
|||||
```

THIS PAGE BLANK (USPTO)

```

RESULT 11
:08218
:isopenicillin N synthase (EC 1.14.11.-) pcbC [similarity] - Flavobacterium sp. (strain S
:Species: Flavobacterium sp.
:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 26-May-2000
:Accession: S08218
:Shiffman, D.; Cohen, G.; Aharonowitz, Y.; Palissa, H.; von Dohren, H.; Kleinkauf, H.;
:Title: Nucleotide sequence of the isopenicillin N synthase gene (pcbC) of the gram neg
:Reference number: S08218; MUID:90174998; PMID:2308852
:Accession: S08218
:Molecule type: DNA
:Residues: 1-326 <SHI>
:CROSS-references: EMBL:X17355; NID:g43423; PIDN:CAA35233.1; PID:g43424
:Genetics:
:Gene: pcbC
:Superfamily: isopenicillin N synthase
:Keywords: antibiotic biosynthesis; ascorbic acid; iron; metalloprotein; oxidoreductase
:45,209,265/Binding site: iron (His) #status predicted

Query Match      9.9%; Score 162; DB 2; Length 326;
Best Local Similarity 22.9%; Pred. No. 6.6e-07;
Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;

iy 24 RCLRDKGLFLYDCGLTDTLTKSAKDLVIDFFEHGSEAEK-----RAVTSPTVTMRRGFT 78
ib 32 RACRGSGFFYAANGV---DLAALQKFTTDWHMAMSAEEKWELAIRAYNPANPRNRNGY- 87
iy 79 GLESESTAQITNTGSDYSCYMGMTAD-----NLFPs-----GDFGRIW 119
ib 88 -----YMAVEGKKANESFCYLNPSFDADHATIKAGLPSHEVNIWPDEARHPGMRFF 139
iy 120 TOYEDRQYTSRAVAREVLRAATGTEPDGGVEAF-----LDCEPLLRFRYPPOVPE 169
ib 140 EAYFSDFVDAVAILRGFAIALGRE-----ESFFERHFSMDDTLSAVSLRYPFLENYP- 193
iy 170 HRSABEQLRMAP-----HYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPD 220
ib 194 -----PLKLGPDGKLSFEHHQDVSLLITVLYQTAPN---LQVETAEGLDIPVSDE 242
iy 221 AVLVFCAIATLVGGQVAPRHVAAAPRRDQIAGSSRTSSVFF 264
ib 243 HFLVNCGTMAHITNGYYPAPVHRV-----KYNAERLSIPFF 280

:Genetics:
:Gene: pcbC
:Superfamily: isopenicillin N synthase
:Keywords: antibiotic biosynthesis; ascorbic acid; iron; metalloprotein; oxidoreductase
:45,209,265/Binding site: iron (His) #status predicted

Query Match      9.9%; Score 162; DB 2; Length 326;
Best Local Similarity 22.9%; Pred. No. 6.6e-07;
Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;

iy 24 RCLRDKGLFLYDCGLTDTLTKSAKDLVIDFFEHGSEAEK-----RAVTSPTVTMRRGFT 78
ib 32 RACRGSGFFYAANGV---DLAALQKFTTDWHMAMSAEEKWELAIRAYNPANPRNRNGY- 87
iy 79 GLESESTAQITNTGSDYSCYMGMTAD-----NLFPs-----GDFGRIW 119
ib 88 -----YMAVEGKKANESFCYLNPSFDADHATIKAGLPSHEVNIWPDEARHPGMRFF 139
iy 120 TOYEDRQYTSRAVAREVLRAATGTEPDGGVEAF-----LDCEPLLRFRYPPOVPE 169
ib 140 EAYFSDFVDAVAILRGFAIALGRE-----ESFFERHFSMDDTLSAVSLRYPFLENYP- 193
iy 170 HRSABEQLRMAP-----HYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPD 220
ib 194 -----PLKLGPDGKLSFEHHQDVSLLITVLYQTAPN---LQVETAEGLDIPVSDE 242
iy 221 AVLVFCAIATLVGGQVAPRHVAAAPRRDQIAGSSRTSSVFF 264
ib 243 HFLVNCGTMAHITNGYYPAPVHRV-----KYNAERLSIPFF 280

:Genetics:
:Gene: pcbC
:Superfamily: isopenicillin N synthase
:Keywords: antibiotic biosynthesis; ascorbic acid; iron; metalloprotein; oxidoreductase
:45,209,265/Binding site: iron (His) #status predicted

Query Match      9.9%; Score 162; DB 2; Length 326;
Best Local Similarity 22.9%; Pred. No. 6.6e-07;
Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;

iy 24 RCLRDKGLFLYDCGLTDTLTKSAKDLVIDFFEHGSEAEK-----RAVTSPTVTMRRGFT 78
ib 32 RACRGSGFFYAANGV---DLAALQKFTTDWHMAMSAEEKWELAIRAYNPANPRNRNGY- 87

```

```

Qy 79 GLESESTAQITNTGSDYSCYMGMTAD-----NLFPs-----GDFGRIW 119
Db 88 -----YMAVEGKKANESFCYLNPSFDADHATIKAGLPSHEVNIWPDEARHPGMRFF 139
Qy 120 TOYEDRQYTSRAVAREVLRAATGTEPDGGVEAF-----LDCEPLLRFRYPPOVPE 169
Db 140 EAYFSDFVDAVAILRGFAIALGRE-----ESFFERHFSMDDTLSAVSLRYPFLENYP- 193
Qy 170 HRSABEQLRMAP-----HYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPD 220
Db 194 -----PLKLGPDGKLSFEHHQDVSLLITVLYQTAPN---LQVETAEGLDIPVSDE 242
Qy 221 AVLVFCAIATLVGGQVAPRHVAAAPRRDQIAGSSRTSSVFF 264
Db 243 HFLVNCGTMAHITNGYYPAPVHRV-----KYNAERLSIPFF 280

RESULT 13
D84713
probable dioxigenase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: D84713
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84713
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <STO>
A:CROSS-references: GB:AE002093; NID:g3201612; PIDN:AAC20719.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g30840
A:Map position: 2
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match      8.9%; Score 146.5; DB 2; Length 362;
Best Local Similarity 23.2%; Pred. No. 1.9e-05;
Matches 68; Conservative 40; Mismatches 118; Indels 67; Gaps 11;

Qy 5 VPTFSLAELOQLHQDEF-----RRCLRDKGLFLYDCGLTDTLTKSAKDLVID 53
Db 57 IPTIDL-----KGVFDEYTVTRESVIAIRDAVERFCFFQVINHGINSNDVMEKMGDGR 112
Qy 54 FFEHSGAEKRAVTSPTVTMRRGFTGLESESTAQITNTGSDYSCYMGMTAD-----N 108
Db 113 FHEQSDVRKKFY-----TRDVTKTYSNFDLYSSPSANWRDLS 154
Qy 109 LFPSGDF-----GRIWTOYFDROYTASRAVAREVLRAATGTEPDGGVEAFDCEP- 157
Db 155 CFMAPDVPETEDLPDICEIMLEAKRMKLGELIFELLSALGNPNHLKE--WDCTYK 212
Qy 158 -LLAFRYPPOVPEHRSABEQLRMAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLP 216
Db 213 LLMLSHYPPCFE-----PGLTFTGTPSHSDRSFLILQ----DHIGGLQVQNGVWDVP 264
Qy 217 YRPDAVLVFCGAIATLVGGQVAPRHVAAAPRRDQIAGSSRTSSV--FFLRP 267
Db 265 PVPGALLVNLGLLQLMTNDQFVSVEHRVLANK-----GEKPRISVASFFVHP 312

RESULT 14
D96635
probable gibberellin 20-oxidase T7P1.12 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: D96635
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.;

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:33:57 ; Search time 42 Seconds
(without alignments)
1175.332 Million cell updates/sec

Title: US-09-582-486-1

Perfect score: 1637

Sequence: 1 MDTTPTFSLAELOQLHQD.....ATFQDWIGGNYNIRRTSKA 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1637	100.0	311	20	AAW22220
2	1629	99.5	311	18	AAW17775
3	1629	99.5	311	23	AAU11044
4	1627	99.4	311	23	AAU11050
5	1626	99.3	311	23	AAU11046
6	1626	99.3	311	23	AAU11051
7	1625	99.3	310	23	AAU11053
8	1625	99.3	310	23	AAU11525
9	1625	99.3	311	18	AAW34168

Expandase mutant R
Expandase mutant R
Streptomyces clavu
Streptomyces clavu
S. clavuligerus mo
Expandase mutant R
Expandase mutant R
Streptomyces clavu
S. clavuligerus pe
Streptomyces clavu
Deacetoxycephaloop
Expandase mutant R
Expandase mutant R
Expandase mutant P
S. clavuligerus ex
Streptomyces clavu
S. clavuligerus pe
S. clavuligerus pe
S. clavuligerus pe
S. clavuligerus pe
S. clavuligerus pe
Expandase mutant C
Expandase mutant P
Expandase mutant C
Expandase mutant C
S. clavuligerus pe
S. clavuligerus pe
S. clavuligerus ex
S. clavuligerus ex
S. clavuligerus ex
S. clavuligerus ex
S. clavuligerus pe
S. clavuligerus ex

ALIGNMENTS

RESULT 1
AAW22220
ID AAY22220 standard; protein; 311 AA.

XX AC AAY22220;
XX DT 16-SEP-1999 (first entry)
XX DE Deacetoxycephalosporin C synthase.

XX KW DAOCS; deacetoxycephalosporin C synthase; antibacterial production;
XX KW three-dimensional crystal structure; beta-lactam production; penicillin;
XX KW cephalosporin; cepham; 2-oxoglutarate dependent enzyme; inhibitor;
XX KW fibrotic disease; liver cirrhosis; arthritis; genetic disorder; therapy;
XX KW herbicide resistance.

XX OS Streptomyces clavuligerus.

XX PN WO9933994-A1.

XX PD 08-JUL-1999.

XX PF 24-DEC-1998; 98WO-GB03860.

XX PR 24-JUN-1998; 98GB-0013644.

XX PR 24-DEC-1997; 97GB-0027370.

XX PA (ISIS-) ISIS INNOVATION LTD.

XX PI Andersson I, Baldwin JE, Hajdu J, Harlos K, Lloyd MD;

PI Ramaswamy S, Roach PL, Schofield CJ, Terwisscha Van Scheltinga AS;

PI Valegard K;

XX WPI; 1999-405515/34.
XX Three-dimensional crystal structures of deacetoxycephalosporin C
PT synthase useful for the production of beta-lactams
XX
XX Claim 1; Page 11; 178pp; English.
XX
XX The invention relates to three-dimensional crystal structures of
CC deacetoxycephalosporin C (DAOCS), which has this amino
CC acid sequence. Micro-organisms capable of expressing a gene encoding a
CC modified DAOCS are useful for the production of beta-lactams of the
CC penicillin or cephalosporin (including cepham) families. The 3D structure
CC of DAOCS is used to determine or predict the structure of another related
CC 2-oxoglutarate dependent enzyme (such as deacetylcephalosporin C synthase
CC (DAOCS)) or related enzymes not from the penicillin and cephalosporin
CC biosynthesis pathway. The structural data is used to modify the other
CC enzyme or for designing an inhibitor for the other enzyme. The other
CC enzyme may be prolyl 4/3 or aspartyl/lysyl-hydroxylase and the inhibitor
CC is used for the treatment of human disease including fibrotic diseases
CC including liver cirrhosis and arthritis. The other enzyme may be
CC p-hydroxyphenylpyruvate dioxygenase and the inhibitor used to treat
CC certain genetic disorders. Alternatively, the other enzyme is involved in
CC herbicide resistance and the information is used to design new herbicides
CC to overcome the problem of resistance. The crystal structure data allows
CC modification of the enzyme to accept (or accept more efficiently)
CC unnatural penicillin substrates for the preparation of new or
CC commercially valuable antibacterial materials, or to enable the modified
CC enzyme to produce unnatural (e.g. exomethylene cepham) or optimise the
CC production of minor products (e.g. 3-beta-hydroxycephams) for use as
CC antibacterials or as intermediates in the preparation of antibacterials
CC or commercially valuable compounds.
XX
XX Sequence 311 AA;
SQ

Query Match 100.0%; Score 1637; DB 20; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.1e-172;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTTPTFSLAELOOGLHODEFRRLDKGLFYLTDGTLDTLTKSAKDLVIDFFHSGE 60
DB 1 MDTTPTFSLAELOOGLHODEFRRLDKGLFYLTDGTLDTLTKSAKDLVIDFFHSGE 60
QY 61 AEKRAVTSVPVPTWRRGFTGLESESTAIQITNTGSDYSYMCYSMTADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVPTWRRGFTGLESESTAIQITNTGSDYSYMCYSMTADNLPSPGDFGRITW 120
QY 121 QYFDRQYTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
DB 121 QYFDRQYTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGATATLVGGQVKA 240
DB 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGATATLVGGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311

RESULT 2
AAW17775
ID AAW17775 standard; protein; 311 AA.

XX AAW17775;
AC AAW17775;
DT 17-FEB-1998 (first entry)
XX
DE S. clavuligerus wild type expandase.

XX Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;
KW thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;
KW semi-synthetic cephalosporin production; penicillin G; isopenicillin N;
KW cephalosporin ring; DOAC.
XX
OS Streptomyces clavuligerus.
XX
PN W09720053-A2.
XX
PD 05-JUN-1997.
XX
PF 27-NOV-1996; 96WO-BP05358.
XX
PR 27-NOV-1995; 95US-0007564.
XX
PR 27-NOV-1995; 95EP-0203259.
XX
PA (KONN) GIST-BROCADES BV.
XX (UYOX-) UNIV OXFORD.
XX
PI Bovenberg RAL, Sutherland JD, Van Der Laan JM;
XX
XX WPI; 1997-310608/28.
XX
XX Expandase enzymes with mutation(s) to alter substrate specificity -
PT used particularly for production of
PT 7-amino-desacetoxycephalosporanic acid
XX
XX Disclosure; Fig 1; 21pp; English.
XX
XX This sequence represents the wild type expandase from Streptomyces
CC clavuligerus. The DNA encoding this sequence was subjected to site
CC directed mutagenesis to create the mutant expandase sequences of the
CC invention (see AAW34166-W34177). The mutant expandase sequences have an
CC altered substrate specificity compared to the wild type expandase.
CC Expandase catalyses the expansion of the 5-membered thiazolidine ring of
CC penicillin N to the 6-membered dihydrothiazine ring of DOAC. The mutant
CC expandase gene is used particularly for the production of
CC 7-aminodesacetoxycephalosporanic acid (7-ADCA). 7-ADCA is one of the key
CC intermediates that is used in the preparation of semi-synthetic
CC cephalosporins. The mutant expandases have improved ability to expand
CC penicillin G into a cephalosporin ring in vitro and/or in vivo where
CC other penicillins such as penicillin N and isopenicillin N can act as
CC competing substrates.
XX
XX Sequence 311 AA;
SQ

Query Match 99.5%; Score 1629; DB 18; Length 311;
Best Local Similarity 99.7%; Pred. No. 8.5e-172;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDTTPTFSLAELOOGLHODEFRRLDKGLFYLTDGTLDTLTKSAKDLVIDFFHSGE 60
DB 1 MDTTPTFSLAELOOGLHODEFRRLDKGLFYLTDGTLDTLTKSAKDLVIDFFHSGE 60
QY 61 AEKRAVTSVPVPTWRRGFTGLESESTAIQITNTGSDYSYMCYSMTADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVPTWRRGFTGLESESTAIQITNTGSDYSYMCYSMTADNLPSPGDFGRITW 120
QY 121 QYFDRQYTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
DB 121 QYFDRQYTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGATATLVGGQVKA 240
DB 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGATATLVGGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311

```

Db      301 NYNIRRTSKA 311
RESULT 3
AAU11044
ID      AAU11044 standard; Protein; 311 AA.
AC      AAU11044;
XX      12-MAR-2002 (first entry)
XX      Streptomyces clavuligerus penicillin N expandase protein sequence.
XX      Penicillin N expandase; ring-expanding; penicillin G;
KW      phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid;
KW      cephalosporin; cephalixin; penicillin V.
XX      Streptomyces clavuligerus.
XX      WO200185951-A1.
XX      15-NOV-2001.
XX      09-MAY-2001; 2001WO-GB02047.
XX      09-MAY-2000; 2000GB-0011185.
XX      (ACSD-) ACS DOBFAR UK LTD.
XX      Johnson RI, Newbert RW;
XX      WPI; 2002-075247/10.
XX      N-PSDB; AAS17241.
XX      New enzyme useful for ring-expanding penicillin G to produce
PT      phenylacetyl-7-amino desacetoxyccephalosporanic acid, comprises
PT      penicillin expandase having increased specificity for substrates such
PT      as penicillin G -
XX      Claim 4; Fig 1; 42pp; English.
XX      The present invention relates to new penicillin expandases modified to
CC      improve the ring-expanding activity for a substrate which is not
CC      the natural substrate of the unmodified expandase. The invention is
CC      useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
CC      (amino desacetoxyccephalosporanic acid) and the phenylacetyl side chain
CC      is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
CC      as a starting point for the production of a range of semi-synthetic
CC      cephalosporins, most notably cephalixin. The polynucleotide of the
CC      invention is useful for producing a primer e.g. a PCR primer and as
CC      a probe. The modification of the penicillin expandase enhances the
CC      activity of the enzyme such as penicillin N expandase for penicillin
CC      G or V as a substrate. The modified expandase has enhanced catalytic
CC      activity or increased specificity for another substrate such as
CC      penicillin G. The present protein sequence represents the penicillin N
CC      expandase enzyme of the invention.
XX      SQ      Sequence      311 AA;
Query Match      99.5%; Score 1629; DB 23; Length 311;
Best Local Similarity 99.74; Fred. No. 8.5e-172;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1      MDTTPTFSLAEQQLHQDEFRCLRDKGLFVLTDCGLDTELKSKDLVDPFRHGS 60
Db      1      MDTTPTFSLAEQQLHQDEFRCLRDKGLFVLTDCGLDTELKSKDLVDPFRHGS 60
QY      61      AEKRAVTSPTVPTWRGFTGLESESTAQITNTGYSYDYSMCYSMTADNLPFGDGRIT 120
Db      61      AEKRAVTSPTVPTWRGFTGLESESTAQITNTGYSYDYSMCYSMTADNLPFGDGRIT 120
QY      121      QYFDRQYTASRAVAREVLRTATGTEPDGGVEAFIDCEPLLRFYFPQVPEHRSABEOPLRM 180

```

CC Note: The present protein sequence is not shown in the specification
CC but is derived from the Streptomyces clavuligerus wild-type
CC penicillin N expandase sequence given in figure 1 (see AAU11044).
XX
SQ Sequence 311 AA;
Query Match 99.4%; Score 1627; DB 23; Length 311;
Best Local Similarity 99.4%; Pred. No. 1.4e-171;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDTTPTFSLAELOQGLHODEFRRLRDKGLFYLTCGLTDTLTKSAKDLVDFPHGSE 60
DB 1 MDTTPTFSLAELOQGLHODEFRRLRDKGLFYLTCGLTDTLTKSAKDLVDFPHGSE 60
QY 61 AEKRAVTSVPVPMRRGFTGLESESTAIQITNTGSYSDYSCYSGMTADNLPFGDGRITW 120
DB 61 AEKRAVTSVPVPMRRGFTGLESESTAIQITNTGSYSDYSCYSGMTADNLPFGDGRITW 120
QY 121 QYFDQYQTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
DB 121 QYFDQYQTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSMTVLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMTVLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIATLVGGQVKA 240
QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311

RESULT 5
AAU11046
ID AAU11046 standard; Protein; 311 AA.
AC AAU11046;
DT 12-MAR-2002 (first entry)
XX Streptomyces clavuligerus penicillin N expandase L158V mutant.
DE Penicillin N expandase; ring-expanding; penicillin G;
XX phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid;
KW cephalosporin; cephalaxin; penicillin V; mutant; mutein.
XX Streptomyces clavuligerus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 158 /note= "Wild-type Leu substituted by Val"
XX
XX WO200185951-A1
XX
PD 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-GB02047.
XX
XX 09-MAY-2000; 2000GB-0011185.
XX
XX (ACSD-) ACS DOBFAR UK LTD.
XX
XX Johnson RI, Newbert RW;
XX WPI; 2002-075247/10.
XX
XX New enzyme useful for ring-expanding penicillin G to produce
PT phenylacetyl-7-amino desacetoxyccephalosporanic acid, comprises
PT penicillin expandase having increased specificity for substrates such

PT as penicillin G -
XX
PS Claim 6; Page -; 42pp; English.
XX
CC The present invention relates to new penicillin expandases modified to
CC improve the ring-expanding activity for a substrate which is not
CC the natural substrate of the unmodified expandase. The invention is
CC useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
CC (amino desacetoxyccephalosporanic acid) and the phenylacetyl side chain
CC is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
CC as a starting point for the production of a range of semi-synthetic
CC cephalosporins, most notably cephalixin. The polynucleotide of the
CC invention is useful for producing a primer e.g. a PCR primer and as
CC a probe. The modification of the penicillin N expandase enhances the
CC activity of the enzyme such as penicillin N expandase for penicillin
CC G or V as a substrate. The modified expandase has enhanced catalytic
CC activity or increased specificity for another substrate such as
CC penicillin G. This sequence is one of a collection of Streptomyces
CC clavuligerus mutants, AAU11046-AAU11054 and AAU11521-AAU11537, used in
CC the invention to improve the ring-expanding activity of expandase.
CC Note: The present protein sequence is not shown in the specification
CC but is derived from the Streptomyces clavuligerus wild-type
CC penicillin N expandase sequence given in figure 1 (see AAU11044).
XX
SQ Sequence 311 AA;
Query Match 99.3%; Score 1626; DB 23; Length 311;
Best Local Similarity 99.4%; Pred. No. 1.8e-171;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDTTPTFSLAELOQGLHODEFRRLRDKGLFYLTCGLTDTLTKSAKDLVDFPHGSE 60
DB 1 MDTTPTFSLAELOQGLHODEFRRLRDKGLFYLTCGLTDTLTKSAKDLVDFPHGSE 60
QY 61 AEKRAVTSVPVPMRRGFTGLESESTAIQITNTGSYSDYSCYSGMTADNLPFGDGRITW 120
DB 61 AEKRAVTSVPVPMRRGFTGLESESTAIQITNTGSYSDYSCYSGMTADNLPFGDGRITW 120
QY 121 QYFDQYQTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
DB 121 QYFDQYQTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSMTVLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMTVLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIATLVGGQVKA 240
QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311
RESULT 6
AAU11051
ID AAU11051 standard; Protein; 311 AA.
XX
XX AAU11051;
XX
DT 12-MAR-2002 (first entry)
XX Streptomyces clavuligerus penicillin N expandase I305M mutant.
DE Penicillin N expandase; ring-expanding; penicillin G;
KW phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid;
KW cephalosporin; cephalixin; penicillin V; mutant; mutein.
XX Streptomyces clavuligerus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH

FT Misc-difference 305 /note= "Wild-type Ile substituted by Met"
FT WO200185951-A1.
FT 15-NOV-2001.
FT 09-MAY-2001; 2001WO-GB02047.
FT 09-MAY-2000; 2000GB-0011185.
FT (ACSD-) ACS DOBPAR UK LTD.
FT Johnson RI, Newbert RW;
FT WPI; 2002-075247/10.
FT New enzyme useful for ring-expanding penicillin G to produce
FT phenylacetyl-7-amino desacetoxyccephalosporanic acid, comprises
FT penicillin N expandase having increased specificity for substrates such
FT as penicillin G -
FT Claim 8; Page -: 42pp; English.
FT
FT The present invention relates to new penicillin expandases modified to
FT improve the ring-expanding activity for a substrate which is not
FT the natural substrate of the unmodified expandase. The invention is
FT useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
FT (amino desacetoxyccephalosporanic acid) and the phenylacetyl side chain
FT is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
FT as a starting point for the production of a range of semi-synthetic
FT cephalosporins, most notably cephalixin. The polynucleotide of the
FT invention is useful for producing a primer e.g. a PCR primer and as
FT a probe. The modification of the penicillin N expandase enhances the
FT activity of the enzyme such as penicillin N expandase for penicillin
FT G or V as a substrate. The modified expandase has enhanced catalytic
FT activity or increased specificity for another substrate such as
FT penicillin G. This sequence is one of a collection of Streptomyces
FT clavuligerus mutants, AAU1046-AAU1054 and AAU1521-AAU1537, used in
FT the invention to improve the ring-expanding activity of expandase.
FT Note: The present protein sequence is not shown in the specification
FT but is derived from the Streptomyces clavuligerus wild-type
FT penicillin N expandase sequence given in figure 1 (see AAU1044).
FT
FT Sequence 311 AA;
FT
FT Query Match 99.3%; Score 1626; DB 23; Length 311;
FT Best Local Similarity 99.4%; Pred. No. 1.8e-171;
FT Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
FT
QY 1 MDTTPTFSLAELOOGLHQDEFRCLRDKGLFYLTCGLTDTLTKSAKDLVIDFFHGE 60
DB 1 MDTTPTFSLAELOOGLHQDEFRCLRDKGLFYLTCGLTDTLTKSAKDLVIDFFHGE 60
QY 61 AEKRAVTSPTVPTMRGFTGLESESTAQITNTGYSYDYSYMGTSMTADNLPSPGDFGRIT 120
DB 61 AEKRAVTSPTVPTMRGFTGLESESTAQITNTGYSYDYSYMGTSMTADNLPSPGDFGRIT 120
QY 121 QYFDROYTASRAVAREVLRTATGTEPDGVEAFIDCEPFLFRFPQVPEHRSABEOPLRM 180
DB 121 QYFDROYTASRAVAREVLRTATGTEPDGVEAFIDCEPFLFRFPQVPEHRSABEOPLRM 180
QY 181 APHYDLSMWTLIOOTPCANGFVSLQAEVGGAFDLPVRPAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMWTLIOOTPCANGFVSLQAEVGGAFDLPVRPAVLVFCGAIATLVGGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFLFRPNADFTFSVPLARECGFVSLDGTATFDQWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFLFRPNADFTFSVPLARECGFVSLDGTATFDQWIGG 300
QY 301 NYNVRTSKA 311
DB 301 NYNVRTSKA 311

RESULT 7
AAU11053
ID AAU11053 standard; Protein; 310 AA.
XX
AC AAU11053;
XX
DT 12-MAR-2002 (first entry)
XX
DE S. clavuligerus penicillin N expandase 1 residue deletion at C-terminus.
XX
KW Penicillin N expandase; ring-expanding; penicillin G;
KW phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid;
KW cephalosporin; cephalixin; penicillin V; mutant; mutein.
XX
OS Streptomyces clavuligerus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 310
FT /note= "Protein is truncated at this residue"
XX
PN WO200185951-A1.
XX
PD 15-NOV-2001.
XX
PF 09-MAY-2001; 2001WO-GB02047.
XX
PR 09-MAY-2000; 2000GB-0011185.
XX
PA (ACSD-) ACS DOBPAR UK LTD.
XX
PI Johnson RI, Newbert RW;
XX
DR WPI; 2002-075247/10.
XX
PT New enzyme useful for ring-expanding penicillin G to produce
PT phenylacetyl-7-amino desacetoxyccephalosporanic acid, comprises
PT penicillin N expandase having increased specificity for substrates such
PT as penicillin G -
XX
PS Disclosure; Page -: 42pp; English.
XX
CC The present invention relates to new penicillin expandases modified to
CC improve the ring-expanding activity for a substrate which is not
CC the natural substrate of the unmodified expandase. The invention is
CC useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
CC (amino desacetoxyccephalosporanic acid) and the phenylacetyl side chain
CC is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
CC as a starting point for the production of a range of semi-synthetic
CC cephalosporins, most notably cephalixin. The polynucleotide of the
CC invention is useful for producing a primer e.g. a PCR primer and as
CC a probe. The modification of the penicillin N expandase enhances the
CC activity of the enzyme such as penicillin N expandase for penicillin
CC G or V as a substrate. The modified expandase has enhanced catalytic
CC activity or increased specificity for another substrate such as
CC penicillin G. This sequence is one of a collection of Streptomyces
CC clavuligerus mutants, AAU1046-AAU1054 and AAU1521-AAU1537, used in
CC the invention to improve the ring-expanding activity of expandase.
CC Note: The present protein sequence is not shown in the specification
CC but is derived from the Streptomyces clavuligerus wild-type
CC penicillin N expandase sequence given in figure 1 (see AAU1044).
CC
CC Sequence 310 AA;
CC
CC Query Match 99.3%; Score 1625; DB 23; Length 310;
CC Best Local Similarity 99.7%; Pred. No. 2.3e-171;
CC Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
QY 1 MDTTPTFSLAELOOGLHQDEFRCLRDKGLFYLTCGLTDTLTKSAKDLVIDFFHGE 60
DB 1 MDTTPTFSLAELOOGLHQDEFRCLRDKGLFYLTCGLTDTLTKSAKDLVIDFFHGE 60

QY 61 AEKRAVTSVPPTMRGFTGLESTAQITNTGSDYSYSCYSGMTADNLPSPGDFGRITW 120
 DB 61 AEKRAVTSVPPTMRGFTGLESTAQITNTGSDYSYSCYSGMTADNLPSPGDFGRITW 120
 QY 121 QYDFDQYTTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
 DB 121 QYDFDQYTTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
 QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKA 240
 DB 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKA 240
 QY 241 PRHHVAAPRRDQTAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 DB 241 PRHHVAAPRRDQTAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 QY 301 NYVNIIRRTSK 310
 DB 301 NYVNIIRRTSK 310

RESULT 8

AAU11525
 ID AAU11525 standard; Protein; 310 AA.

XX AC AAU11525;
 XX DT 12-MAR-2002 (first entry)
 XX DE S. clavuligerus penicillin N expandase K310 deletion mutant.
 XX KW Penicillin N expandase; ring-expanding; penicillin G;
 KW phenylacetyl-7-ADCA; amino desacetoxycephalosporanic acid;
 XX cephalosporin; cephalixin; penicillin V; mutant; mutein.
 OS Streptomycetes clavuligerus.
 OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 310 /note= "Protein is truncated at this residue"

XX PN WO200185951-A1.

XX PD 15-NOV-2001.

XX PF 09-MAY-2001; 2001WO-GB02047.

XX PR 09-MAY-2000; 2000GB-0011185.

XX PA (ACSD-) ACS DOBFAR UK LTD.

XX PI Johnson RI, Newbert RW;

XX DR WPI; 2002-075247/10.

XX PT New enzyme useful for ring-expanding penicillin G to produce
 PT phenylacetyl-7-amino desacetoxycephalosporanic acid, comprises
 PT penicillin expandase having increased specificity for substrates such
 as penicillin G -

XX PS Examples; Page -; 42pp; English.

XX CC The present invention relates to new penicillin expandases modified to
 CC improve the ring-expanding activity for a substrate which is not
 CC the natural substrate of the unmodified expandase. The invention is
 CC useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
 CC (amino desacetoxycephalosporanic acid) and the phenylacetyl side chain
 CC is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
 CC as a starting point for the production of a range of semi-synthetic
 CC cephalosporins, most notably cephalixin. The polynucleotide of the
 CC invention is useful for producing a primer e.g. a PCR primer and as

CC a probe. The modification of the penicillin expandase enhances the
 CC activity of the enzyme such as penicillin N expandase for penicillin
 CC G or V as a substrate. The modified expandase has enhanced catalytic
 CC activity or increased specificity for another substrate such as
 CC penicillin G. This sequence is one of a collection of Streptomycetes
 CC clavuligerus mutants, AAU11046-AAU11054 and AAU11521-AAU11537, used in
 CC the invention to improve the ring-expanding activity of expandase.
 CC Note: The present protein sequence is not shown in the specification
 CC but is derived from the Streptomycetes clavuligerus wild-type
 CC penicillin N expandase sequence given in figure 1 (see AAU11044).

XX SQ Sequence 310 AA;

Query Match 99.3%; Score 1625; DB 23; Length 310;
 Best Local Similarity 99.7%; Pred. No. 2.3e-171;
 Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTTPTFSLAEQLQGLHODEFRRCRLDKGLFVLTDCGLTDTLKSADLVIDFFEHGSE 60

DB 1 MDTTPTFSLAEQLQGLHODEFRRCRLDKGLFVLTDCGLTDTLKSADLVIDFFEHGSE 60

QY 61 AEKRAVTSVPPTMRGFTGLESTAQITNTGSDYSYSCYSGMTADNLPSPGDFGRITW 120

DB 61 AEKRAVTSVPPTMRGFTGLESTAQITNTGSDYSYSCYSGMTADNLPSPGDFGRITW 120

QY 121 QYDFDQYTTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180

DB 121 QYDFDQYTTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180

QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKA 240

DB 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKA 240

QY 241 PRHHVAAPRRDQTAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300

DB 241 PRHHVAAPRRDQTAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300

QY 301 NYVNIIRRTSK 310

DB 301 NYVNIIRRTSK 310

RESULT 9

AAW34168

ID AAW34168 standard; protein; 311 AA.

XX AC AAW34168;

XX DT 17-FEB-1998 (first entry)

XX DE Expandase mutant R74Q.

XX KW Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;
 KW thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;
 KW semi-synthetic cephalosporin production; penicillin G; isopenicillin N;
 KW cephalosporin ring; DOAC; mutant.

XX OS Streptomycetes clavuligerus.

OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 74 /note= "Arg to Gln mutation"

XX PN WO9720053-A2.

XX PD 05-JUN-1997.

XX PF 27-NOV-1996; 96WO-EP05358.

XX PR 27-NOV-1995; 95US-0007564.

XX PR 27-NOV-1995; 95EP-0203259.

QY 301 NYVNIIRRTSKA 311
 DB 301 NYVNIIRRTSKA 311

RESULT 11
 AAW34175
 ID AAW34175 standard; protein; 311 AA.
 XX
 AC AAW34175;
 XX
 DT 17-FEB-1998 (first entry)
 XX
 DE Expandase mutant R266N.
 XX
 KW Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;
 KW thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;
 KW semi-synthetic cephalosporin production; penicillin G; isopenicillin N;
 KW cephalosporin ring; DOAC; mutant.
 XX
 OS Streptomyces clavuligerus.
 OS Synthetic.

Key Location/Qualifiers
 FH Misc-difference 266
 FT /note= "Arg to Asn mutation"
 FT
 PN WO9720053-A2.
 XX
 XX 05-JUN-1997.
 XX
 XX 27-NOV-1996; 96WO-EF05358.
 XX
 XX 27-NOV-1995; 95US-0007564.
 PR 27-NOV-1995; 95EP-0203259.
 XX
 XX (KONN) GIST-BROCADES BV.
 PA (UYOX-) UNIV OXFORD.
 XX
 PI Bovenberg RAL, Sutherland JD, Van Der Laan JM;
 XX
 XX WPI; 1997-310608/28.
 DR
 XX
 XX Expandase enzymes with mutation(s) to alter substrate specificity -
 PT used particularly for production of
 PT 7-amino:desacetoxy:cephalosporanic acid
 XX
 XX +Claim 2; Page -: 21pp; English.

AAW34166-W34177 represent mutant expandase sequences of the invention.
 CC These sequences represent mutated versions of the Streptomyces
 CC clavuligerus expandase sequences shown in AAW17775. The DNA encoding the
 CC wild type expandase was subjected to site directed mutagenesis to create
 CC these sequences. The mutant expandase sequences have an altered
 CC substrate specificity compared to the wild type expandase. Expandase
 CC catalyses the expansion of the 5-membered thiazolidine ring of penicillin
 CC N to the 6-membered dihydrothiazine ring of DOAC. The mutant expandase
 CC gene is used particularly for the production of
 CC 7-aminodesacetoxycephalosporanic acid (7-ADCA). 7-ADCA is one of the key
 CC intermediates that is used in the preparation of semi-synthetic
 CC cephalosporins. These sequences have improved ability to expand
 CC penicillin G into a cephalosporin ring in vitro and/or in vivo where
 CC other penicillins such as penicillin N and isopenicillin N can act as
 CC competing substrates.

XX
 SQ Sequence 311 AA;
 Query Match 99.2%; Score 1624; DB 18; Length 311;
 Best Local Similarity 99.4%; Pred No. 3e-171;
 Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDTTPTFSLAELOOGLHQBDFRRCRLDKGLFYLTDGCLTDTTELKSAKDVLIDFFEHGSE 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 1 MDTTPTFSLAELOOGLHQBDFRRCRLDKGLFYLTDGCLTDTTELKSAKDVLIDFFEHGSE 60
 QY 61 AEKRAVTSFVPTMRGFTGLESESTAOITNTGYSYDYSNCSMGTAADNLFFPSGDFGRWT 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 AEKRAVTSFVPTMRGFTGLESESTAOITNTGYSYDYSNCSMGTAADNLFFPSGDFGRWT 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 QYFDRQYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYFFQVPEHRSABEQPLRM 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 121 QYFDRQYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYFFQVPEHRSABEQPLRM 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 APHYDLISWVTLIQOTPCANGFVSLQAEVGGATDLPYRPDAVLVFCGAIATLVGGQVKA 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 181 APHYDLISWVTLIQOTPCANGFVSLQAEVGGATDLPYRPDAVLVFCGAIATLVGGQVKA 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 PRHHVAAPRRDQIAGSSRTSSVFRLPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 241 PRHHVAAPRRDQIAGSSRTSSVFRLPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 301 NYVNIIRRTSKA 311
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 301 NYVNIIRRTSKA 311
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 12
 AAE12016
 ID AAE12016 standard; Protein; 311 AA.
 XX
 AC AAE12016;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Streptomyces clavuligerus modified deacetoxycephalosporin C synthetase.
 XX
 KW Deacetoxycephalosporin C synthetase; DAOCS; antibiotic; penicillin;
 KW oxygenase; non-haeme iron (II) dependent family; oxidase; mutant; muten.
 XX
 OS Streptomyces clavuligerus.
 OS Synthetic.

Key Location/Qualifiers
 FH Misc-difference 181
 FT /note= "Wild type Ala substituted with Arg"
 FT
 PN US6284483-B1.
 XX
 XX 04-SEP-2001.
 XX
 XX 06-OCT-1999; 99US-0413231.
 XX
 XX 06-OCT-1999; 99US-0413231.
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Dilley DR, Kadyrzhanova DK, Wang Z, Warner TW;
 XX
 XX WPI; 2001-615433/71.
 XX
 XX Method for producing antibiotics, particularly penicillin G or V,
 PT comprises employing a modified isopenicillin N synthetase, in either
 PT an organism or a cell-free system under the control of bicarbonate -
 XX
 PS Disclosure; Column 45-48; 27pp; English.
 XX
 CC The invention relates to a modified enzyme of a non-haeme iron (II)
 CC dependent family of oxygenases and oxidases which renders the enzyme
 CC dependent on bicarbonate for activity. The invention also related to a
 CC method for producing penicillin G or V comprises employing a modified
 CC enzyme, particularly isopenicillin N synthetase (IPNS), in either an
 CC organism or a cell-free system, under the control of bicarbonate. The
 CC modified enzymes are IPNS, deacetoxycephalosporin C synthetase (DAOCS)
 CC and deacetoxycephalosporin C synthetase (DACS). The method is useful for
 CC producing antibiotics, particularly penicillin G or V. The method is
 CC useful for making organisms useful for making an antibiotic dependent

CC on bicarbonate to make the antibiotic. The present sequence is
CC Streptomyces clavuligerus modified deacetoxycephalosporin C synthetase
CC showing an Arg181 in place of Ala181.
XX
SQ Sequence 311 AA;

Query Match 99.2%; Score 1624; DB 22; Length 311;
Best Local Similarity 99.4%; Pred. No. 3e-171;
Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDTTPTFSLAEHQGLHQBDEFRCLRDKGLFYLDCGLTDTLTKSAKOLVIDFFEHGSE 60
DB 1 MDTTPTFSLAEHQGLHQBDEFRCLRDKGLFYLDCGLTDTLTKSAKOLVIDFFEHGSE 60

QY 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120

QY 121 QYFDRQYTSRAVAREVLRTATGTEPDGGEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
DB 121 QYFDRQYTSRAVAREVLRTATGTEPDGGEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180

QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGOVKA 240
DB 181 RPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGOVKA 240

QY 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300

QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311

RESULT 13
AAU11047
ID AAU11047 standard; Protein; 311 AA.
XX
AC AAU11047;
XX
DT 12-MAR-2002 (first entry)
XX
DE Streptomyces clavuligerus penicillin N expandase L158A mutant.
XX
KW Penicillin N expandase; ring-expanding; penicillin G;
KW phenylacetyl-7-ADCA; amino desacetoxycephalosporanic acid;
KW cephalosporin; cephalaxin; penicillin V; mutant; muten.
XX
OS Streptomyces clavuligerus.
OS Synthetic.
XX
PH Key Location/Qualifiers
FT Misc-difference 158 /note= "Wild-type Leu substituted by Ala"
FT
PN WO200185951-A1.
XX
PD 15-NOV-2001.
XX
PF 09-MAY-2001; 2001WO-GB02047.
XX
PR 09-MAY-2000; 2000GB-0011185.
XX
PA (ACSD-) ACS DOBFAR UK LTD.
XX
PI Johnson RI, Newbert RW;
XX
DR WPI; 2002-075247/10.
XX
PT New enzyme useful for ring-expanding penicillin G to produce
PT phenylacetyl-7-amino desacetoxycephalosporanic acid, comprises
PT penicillin N expandase having increased specificity for substrates such

as penicillin G -
Examples; Page -: 42pp; English.

The present invention relates to new penicillin expandases modified to
improve the ring-expanding activity for a substrate which is not
the natural substrate of the unmodified expandase. The invention is
useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
(amino desacetoxycephalosporanic acid) and the phenylacetyl side chain
is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
as a starting point for the production of a range of semi-synthetic
cephalosporins, most notably cephalaxin. The polynucleotide of the
invention is useful for producing a primer e.g. a PCR primer and as
a probe. The modification of the penicillin expandase enhances the
activity of the enzyme such as penicillin N expandase for penicillin
G or V as a substrate. The modified expandase has enhanced catalytic
activity or increased specificity for another substrate such as
penicillin G. This sequence is one of a collection of Streptomyces
clavuligerus mutants, AAU1046-AAU11054 and AAU1521-AAU1537, used in
the invention to improve the ring-expanding activity of expandase.
CC Note: The present protein sequence is not shown in the specification
CC but is derived from the Streptomyces clavuligerus wild-type
CC penicillin N expandase sequence given in figure 1 (see AAU11044).
XX
SQ Sequence 311 AA;

Query Match 99.2%; Score 1624; DB 23; Length 311;
Best Local Similarity 99.4%; Pred. No. 3e-171;
Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDTTPTFSLAEHQGLHQBDEFRCLRDKGLFYLDCGLTDTLTKSAKOLVIDFFEHGSE 60
DB 1 MDTTPTFSLAEHQGLHQBDEFRCLRDKGLFYLDCGLTDTLTKSAKOLVIDFFEHGSE 60

QY 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120

QY 121 QYFDRQYTSRAVAREVLRTATGTEPDGGEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
DB 121 QYFDRQYTSRAVAREVLRTATGTEPDGGEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180

QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGOVKA 240
DB 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGOVKA 240

QY 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300

QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311

RESULT 14
ABG73769
ID ABG73769 standard; Protein; 311 AA.
XX
AC ABG73769;
XX
DT 03-APR-2003 (first entry)
XX
DE S. clavuligerus modified DAOCS SEQ ID 18.
XX
KW Non-heme iron II dependent family; oxygenase; oxidase; antibiotic;
KW iron ligand; enzyme; bicarbonate; penicillin G; penicillin V;
KW cephalosporin C; isopenicillin N synthetase; IPNS; DAOCS;
KW deacetoxycephalosporin C synthetase.
XX
OS Streptomyces clavuligerus.
OS Synthetic.
XX

```
PN US2002127633-A1.
XX
PD 12-SEP-2002.
XX
XX 08-AUG-2001; 2001US-0924841.
XX
PF 06-OCT-1999; 99US-0413231.
XX
XX (UNMS ) UNIV MICHIGAN STATE.
PA
PI Dillee DR, Kadyrzhanova DK, Wang Z, Warner TM;
XX WPI; 2003-147393/14.
DR
XX
XX Novel enzyme of non-heme iron II dependent family of oxygenases and
PT oxidases useful for producing antibiotic, has mutation that renders the
PT enzyme dependent on bicarbonate to produce the antibiotic
XX
XX Claim 5; Page 24-25; 31pp; English.
PS
XX This invention describes a novel enzyme of a non-heme iron II dependent
CC family of oxygenases and oxidases which is in a pathway to produce an
CC antibiotic. The enzyme comprises a mutation which is an amino acid
CC residue that is two amino acid residues upstream of a histidine residue
CC which is an iron ligand of the enzyme, where the mutation renders the
CC enzyme dependent on bicarbonate to produce the antibiotic. The enzyme
CC and modified organisms containing the enzyme are useful for producing an
CC antibiotic such as penicillin G or V and cephalosporin C. This sequence
CC represents an deacetoxycephalosporin C synthetase (DAOCS) variant
CC described in the disclosure of the invention.
XX
XX Sequence 311 AA;
SQ
Query Match 99.2%; Score 1624; DB 24; Length 311;
Best Local Similarity 99.4%; Pred. No. 3e-171;
Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MDTTPTFSLAELOOGLHODEFRRCRLDKGLFVLTDCGLTDTLTKSAKDLVIDFFHSGE 60
DB 1 MDTTPTFSLAELOOGLHODEFRRCRLDKGLFVLTDCGLTDTLTKSAKDLVIDFFHSGE 60
QY 61 AEKRAVTSVPVPTMRGFTGLESESTAQIINTGSDYSYSCYSGMTADNLFPPSGDGRITW 120
DB 61 AEKRAVTSVPVPTMRGFTGLESESTAQIINTGSDYSYSCYSGMTADNLFPPSGDGRITW 120
QY 121 QYFDRQYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFYFPQVPEHRSABEQPLRM 180
DB 121 QYFDRQYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFYFPQVPEHRSABEQPLRM 180
QY 181 APHYDLSMVTLLIQTTCANGFVSLQAEVGAFTDLPYRPDAVLVFCGATATLVGGQVKA 240
DB 181 RPHYDLSMVTLLIQTTCANGFVSLQAEVGAFTDLPYRPDAVLVFCGATATLVGGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYNIRRTSKA 311
DB 301 NYNIRRTSKA 311
RESULT 15
AAW34167
ID AAW34167 standard; protein; 311 AA.
XX
AC AAW34167;
XX
XX 17-FEB-1998 (first entry)
DE
XX Expandase mutant R74M.
XX
XX Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;
KW thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;
KW semi-synthetic cephalosporin production; penicillin G; isopenicillin N;
KW cephalosporin ring; DOAC; mutant.
XX
OS Streptomyces clavuligerus.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 74 /note= "Arg to Met mutation"
FT
XX
XX WO9720053-A2.
XX
XX 05-JUN-1997.
PD
XX
XX 27-NOV-1996; 96WO-EP05358.
PF
XX
XX 27-NOV-1995; 95US-0007564.
PR
XX 27-NOV-1995; 95EP-0203259.
PR
XX (KONN ) GIST-BROCADES BV.
PA (OYOK-) UNIV OXFORD.
PA
XX Bovenberg RAL, Sutherland JD, Van Der Laan JM;
XX WPI; 1997-310608/28.
PI
XX
XX Expandase enzymes with mutation(s) to alter substrate specificity -
PT used particularly for production of
PT 7-amino:desacetoxy:cephalosporanic acid
XX
XX Claim 2; Page -; 21pp; English.
PS
XX
XX AAW34166-W34177 represent mutant expandase sequences of the invention.
CC These sequences represent mutated versions of the Streptomyces
CC clavuligerus expandase sequences shown in AAW17775. The DNA encoding the
CC wild type expandase was subjected to site directed mutagenesis to create
CC these sequences. The mutant expandase sequences have an altered
CC substrate specificity compared to the wild type expandase. Expandase
CC catalyses the expansion of the 5-membered thiazolidine ring of penicillin
CC N to the 6-membered dihydrothiazine ring of DOAC. The mutant expandase
CC gene is used particularly for the production of
CC 7-aminodesacetoxycephalosporanic acid (7-ADCA). 7-ADCA is one of the key
CC intermediates that is used in the preparation of semi-synthetic
CC cephalosporins. These sequences have improved ability to expand
CC penicillin G into a cephalosporin ring in vitro and/or in vivo where
CC other penicillins such as penicillin N and isopenicillin N can act as
CC competing substrates.
XX
XX Sequence 311 AA;
SQ
Query Match 99.1%; Score 1623; DB 18; Length 311;
Best Local Similarity 99.4%; Pred. No. 3.9e-171;
Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MDTTPTFSLAELOOGLHODEFRRCRLDKGLFVLTDCGLTDTLTKSAKDLVIDFFHSGE 60
DB 1 MDTTPTFSLAELOOGLHODEFRRCRLDKGLFVLTDCGLTDTLTKSAKDLVIDFFHSGE 60
QY 61 AEKRAVTSVPVPTMRGFTGLESESTAQIINTGSDYSYSCYSGMTADNLFPPSGDGRITW 120
DB 61 AEKRAVTSVPVPTMRGFTGLESESTAQIINTGSDYSYSCYSGMTADNLFPPSGDGRITW 120
QY 121 QYFDRQYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFYFPQVPEHRSABEQPLRM 180
DB 121 QYFDRQYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFYFPQVPEHRSABEQPLRM 180
QY 181 APHYDLSMVTLLIQTTCANGFVSLQAEVGAFTDLPYRPDAVLVFCGATATLVGGQVKA 240
DB 181 APHYDLSMVTLLIQTTCANGFVSLQAEVGAFTDLPYRPDAVLVFCGATATLVGGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
```

Db 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
Qy 301 NYVNIIRTSKA 311
Db 301 NYVNIIRTSKA 311

Search completed: February 3, 2004, 17:36:36
Job time : 43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:33:58 ; Search time 17 Seconds
(without alignments)
860.312 Million cell updates/sec

Title: US-09-582-486-1
Perfect score: 1637
Sequence: 1 MDTTPTFSLAEIQQHQD.....ATFQDWIGGNNIRTSKA 311

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1629	99.5	311	1	CEFE_STRCL
2	1186.5	72.5	314	1	CEFE_NOCCLA
3	965.5	59.0	310	1	CEFF_NOCCLA
4	939.5	57.4	318	1	CEFF_STRCL
5	926.5	56.6	332	1	EXPA_CEPAC
6	179	10.9	329	1	IPNS_STRJU
7	176	10.8	329	1	IPNS_STRGR
8	164	10.0	329	1	IPNS_STRCL
9	162	9.9	326	1	IPNS_FLASS
10	162	9.9	326	1	IPNS_LYSLA
11	143	8.7	333	1	IPNS_STRLP
12	140	8.6	348	1	FLS_PETHY
13	138	8.4	328	1	IPNS_NOCCLA
14	137.5	8.4	335	1	FLS_CITUN
15	137	8.4	321	1	IPNS_STRCT
16	134.5	8.2	349	1	FLS_SOLTU
17	133.5	8.2	331	1	IPNS_PENCH
18	127	7.8	334	1	FLS_EUSCR
19	122.5	7.5	331	1	IPNS_EMENI
20	121	7.4	356	1	LDOX_ARATH
21	118	7.2	365	1	FL3H_DIACA
22	113.5	6.9	337	1	FLS_MALDO
23	110	6.7	337	1	ISP7_SCHPO
24	109	6.7	336	1	FLS1_ARATH
25	108	6.6	356	1	FL3H_CALCH
26	106	6.5	338	1	IPNS_CEPAC
27	103.5	6.3	395	1	LDOX_MAIZE
28	103.5	6.3	430	1	LDOX_PETHY
29	101	6.2	357	1	LDOX_MALDO
30	99	6.0	358	1	FL3H_ARATH
31	96	5.9	357	1	FL3H_MATIN
32	95	5.9	364	1	FL3H_MALDO
33	95.5	5.8	291	1	FLS_MATIN

RESULT 1				
ID	CEFE_STRCL	STANDARD;	PRT;	311 AA.
AC	P18548;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Deacetoxycephalosporin C synthetase (EC 1.14.20.1) (DAOS)			
DE	(Expandase).			
GN	CEFE.			
OS	Streptomyces clavuligerus.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Streptomycineae; Streptomycetaceae; Streptomyces.			
OX	NCBI_TaxID=1901;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 27064 / DSM 738 / NRRL 3585;			
RX	MEDLINE=89123150; PubMed=2644235;			
RA	Kovacevic S., Weigel B.J., Tobin M.B., Ingolia T.D., Miller J.R.;			
RT	"Cloning, characterization, and expression in Escherichia coli of the			
RT	Streptomyces clavuligerus gene encoding deacetoxycephalosporin C			
RT	synthetase.";			
RL	J. Bacteriol. 171:754-760(1989).			
CC	-I- FUNCTION: Catalyzes the step from penicillin N to deacetoxy-			
CC	cephalosporin C.			
CC	-I- CATALYTIC ACTIVITY: Penicillin N + 2-oxoglutarate + O(2) =			
CC	deacetoxycephalosporin C + succinate + CO(2) + H(2)O.			
CC	-I- COFACTOR: Iron and ascorbate.			
CC	-I- PATHWAY: Cephalosporin antibiotics biosynthesis			
CC	-I- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF			
CC	OXIDOREDUCTASES. STRONG, TO CEFF.			

This SWISS-PROT entry is copyright. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL outstation -				
the European Bioinformatics Institute. There are no restrictions on its				
use by non-profit institutions as long as its content is in no way				
modified and this statement is not removed. Usage by and for commercial				
entities requires a license agreement (See http://www.isb-sib.ch/announce/				
or send an email to license@sib-sib.ch).				

EMBL:	M32324; AAA26715.1; --			
DR	PIR; T52312; T52312.			
DR	PDB; 1DCS; 08-JUN-99.			
DR	PDB; 1ESH; 26-JUN-01.			
DR	PDB; 1ESI; 23-NOV-01.			
DR	PDB; 1HJP; 01-JUN-01.			
DR	PDB; 1HJG; 01-JUN-01.			
DR	PDB; 1RXF; 08-JUN-99.			
DR	PDB; 1RXG; 08-JUN-99.			
DR	InterPro; IPR005123; 2OG-FeII_Oxy.			
DR	InterPro; IPR002057; Isopen_N_synth.			
DR	Pfam; PF031171; 2OG-FeII_Oxy_I.			
DR	PROSITE; PS00185; IPNS_1; FALSE_NEG.			
DR	PROSITE; PS00186; IPNS_2; 1.			
KW	Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C;			
KW	3D-structure.			

P33212 rhizobium f
P55712 rhizobium s
Q9RYV1 deinococcus
P24397 hyoscyamus
P51093 vitis vinif
Q07353 petunia hyb
P10967 lycopersico
P46550 caenorhabdi
P28038 hordeum vul
P41090 vitis vinif
P22248 azotobacter
Q8trk8 methanosarc

ALIGNMENTS

SQ SEQUENCE 311 AA; 34555 MW; 9C64E1FC37F524BC CRC64;
Query Match 99.5%; Score 1629; DB 1; Length 311;
Best Local Similarity 99.7%; Pred. No. 4.4e-143;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDTTVTFSLAELQQGLHQDEFFRCLRDKGLFYLTDGCLTDTLTKSAKDLVIDFFHSGSE 60
DB 1 MDTTVTFSLAELQQGLHQDEFFRCLRDKGLFYLTDGCLTDTLTKSAKDLVIDFFHSGSE 60
QY 61 AEKRAVTSVPVTRRGFTGLSESTAIQITNTGYSYDSCYMGTAADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVTRRGFTGLSESTAIQITNTGYSYDSCYMGTAADNLPSPGDFGRITW 120
QY 121 QYDFROYTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRYPQVPEHRSABEOPLM 180
DB 121 QYDFROYTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRYPQVPEHRSABEOPLM 180
QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGGAFDLPYRDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGGAFDLPYRDAVLVFCGAIATLVGGQVKA 240
QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTSFVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTSFVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYNIRRTSKA 311
DB 301 NYNIRRTSKA 311
RESULT 2
CEFE_NOCCLA STANDARD; PRT; 314 AA.
AC Q03047;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deacetoxycephalosporin C synthetase (EC 1.14.20.1) (DAOCS)
DE (Expandase).
GN CEF.
OS Nocardia lactamdurans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93173127; PubMed=8437592;
RA Coque J.J.R., Martin J.F., Liras P.;
RT "Characterization and expression in Streptomyces lividans of cefD and
RT cefE genes from Nocardia lactamdurans; the organization of the
RT cephamycin gene cluster differs from that in Streptomyces
RT clavuligerus.";
RL Mol. Gen. Genet. 236:453-458(1993).
CC -1- FUNCTION: Catalyzes the step from penicillin N to deacetoxy-
CC cephalosporin C.
CC -1- CATALYTIC ACTIVITY: Penicillin N + 2-oxoglutarate + O(2) =
CC deacetoxycephalosporin C + succinate + CO(2) + H(2)O.
CC -1- COFACTOR: Iron and ascorbate.
CC -1- PATHWAY: Cephalosporin antibiotics biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES. STRONG, TO CEF.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.

CC EMBL; Z13974; CAA78376.1; -;
CC PIR; S30900; S30900.

DR InterPro; IPR005123; 2OG-FelI_Oxy.
DR InterPro; IPR002057; Isopen N synth.
DR Pfam; PF03171; 2OG-FelI_Oxy; 1.
DR PROSITE; PS00185; IPNS_1; FALSE_NEG.
DR PROSITE; PS00186; IPNS_2; FALSE_NEG.
KW Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C.
SQ SEQUENCE 314 AA; 34530 MW; D007AD3DEG7A1BDF CRC64;
Query Match 72.5%; Score 1186.5; DB 1; Length 314;
Best Local Similarity 69.8%; Pred. No. 3.6e-102;
Matches 217; Conservative 41; Mismatches 52; Indels 1; Gaps 1;
QY 2 DTTVTFSLAELQQGLHQDEFFRCLRDKGLFYLTDGCL-TDTLSAKDLVIDFFHSGSE 60
DB 3 DATVTFDLAELREGHQEERFCHLREKGVFYLKGTGLPAEADHASGREIAVDFFDHGTE 62
QY 61 AEKRAVTSVPVTRRGFTGLSESTAIQITNTGYSYDSCYMGTAADNLPSPGDFGRITW 120
DB 63 AEKKAAMTPIPTIRRGYAGLESESTAIQITNTGKYTDYSMSYMGTAADNLPSPAEFEKAW 122
QY 121 QYDFROYTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRYPQVPEHRSABEOPLM 180
DB 123 DYFARMYRASQDVAKQVLTSVGAEPVEGMDAFDCEPLLRFLRYFFPEVDRVAEOPLM 182
QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGGAFDLPYRDAVLVFCGAIATLVGGQVKA 240
DB 183 APHYDLSIVTLIHQTPCANGFVSLQVEVDGSYVDIPAQFCAVLVFCGAVATLVADGA 242
QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTSFVPLARECGFDVSLDGETATFQDWIGG 300
DB 243 PKHVAAPGADKRVGSSRTSSVFFLRPNGDFFRSVPARECGFDVSIPTAETATFDDWIGG 302
QY 301 NYNIRRTSKA 311
DB 303 NYNIRRTSKA 313
RESULT 3
CEFE_NOCCLA STANDARD; PRT; 310 AA.
AC F42219;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deacetoxycephalosporin C hydroxylase (EC 1.14.11.-)
DE (Deacetylcephalosporin C synthetase) (DACS) (Beta-lactam hydroxylase).
GN CEF.
OS Nocardia lactamdurans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96287477; PubMed=8703431;
RA Coque J., Enguita F.J., Cardoza R.E., Martin J.F., Liras P.;
RT "Characterization of the cefD gene of Nocardia lactamdurans encoding
RT a 3'-methylcephem hydroxylase different from the 7-cepem
RT hydroxylase.";
RL Appl. Microbiol. Biotechnol. 44:605-609(1996).
CC -1- FUNCTION: HYDROXYLATION OF DESACETOXYCEPHALOSPORIN C IN
CC 3'-POSITION TO FORM DEACETYLCEPHALOSPORIN C.
CC -1- PATHWAY: Cephalosporin biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES. STRONG, TO CEF.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.

CC EMBL; Z13974; CAA78376.1; -;
CC PIR; S30900; S30900.

```
DR EMBL; Z21687; CAA79803.1; -.
DR PIR; S40253;
DR InterPro; IPR005123; 2OG-Fell_Oxy.
DR Pfam; PF03171; 2OG-Fell_Oxy; 1.
DR Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C.
KW SEQUENCE 310 AA; 34364 MW; 9983EFAC1B41656 CRC64;
SQ
Query Match 59.0%; Score 965.5; DB 1; Length 310;
Best Local Similarity 61.2%; Pred. No. 9.4e-82;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;
QY 2 DTTVPFSLAELOOGLHODEFRCLDKGLFYLTDCLGTLDTLTKSAKOLVIDFPEHGSEA 61
DB 3 DKTVPFSLAELOOGLHODEFRCLDKGLFYLTDCLGTLDTLTKSAKOLVIDFPEHGSEA 61
QY 62 EKRAVTSPTMRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRIWTQ 121
DB 62 EKQAVTKVPTMRRGVSALAEASTAQVNTGTTDYSMSYSGMIGNLPSPKESFESVWTD 121
QY 122 YFDROYTASRAVAREVLRTATGTEPDGGVEAFDCBELLRFRYPQVPEHRSAAEQPLMA 181
DB 122 YFDSLRYAAQETARLVLTAAAGTVGDELDLTLDCDFVLRVLPPEVPEHRAAEYEPERMA 181
QY 182 PHYDLSMVTLIQOTPCANGFVSLOAEVGAFTDLPYRDPDAVLVFCGAIATLVTGGQVAP 241
DB 182 PHYDLSMVTLIQOTPCANGFVSLOAEVGAFTDLPYRDPDAVLVFCGAIATLVTGGQVAP 241
QY 242 RHVAAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVDSLDGETATFQDWI 301
DB 242 NHVSPDASMLAGSDRTSSVFFLRPNADFTFSVPLARECGFVDSLDGETATFQDWI 301
QY 302 YVNI 305
DB 302 YVTM 305
RESULT 4
CEFF_STRCL STANDARD; PRT; 318 AA.
AC P42270;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deacetoxycephalosporin C hydroxylase (EC 1.14.11.-)
DE (Deacetylcephalosporin C synthetase) (DACS) (Beta-lactam hydroxylase).
GN CEFF.
OS Streptomyces clavuligerus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1901;
RN [1]
RP MEDLINE=91100311; PubMed=1987130;
RA Kovacevic S., Miller J.R.;
RT "Cloning and sequencing of the beta-lactam hydroxylase gene (ceff)
RT from Streptomyces clavuligerus: Gene duplication may have led to
RT separate hydroxylase and expandase activities in the actinomycetes.";
RL J. Bacteriol. 173:398-400(1991).
CC
CC -1- FUNCTION: HYDROXYLATION OF DESACETOXICEPHALOSPORIN C IN
CC 3' POSITION TO FORM DEACETYLCEPHALOSPORIN C.
CC -1- PATHWAY: Cephalosporin biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES. STRONG, TO CEFF.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M63809; AAA26716.1; -.

```

```
DR PIR; A39204; A39204.
DR InterPro; IPR005123; 2OG-Fell_Oxy.
DR Pfam; PF03171; 2OG-Fell_Oxy; 1.
DR Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C.
KW SEQUENCE 318 AA; 34584 MW; B17CC1CBC1E67178 CRC64;
SQ
Query Match 57.4%; Score 939.5; DB 1; Length 318;
Best Local Similarity 58.6%; Pred. No. 2.5e-79;
Matches 180; Conservative 43; Mismatches 81; Indels 3; Gaps 1;
QY 2 DTTVPFSLAELOOGLHODEFRCLDKGLFYLTDCLGTLDTLTKSAKOLVIDFPEHGSEA 61
DB 3 DTPVIFNLALREGADQEKRECVTGMGVFLTYGAGDKHRLATDTAMDFFANGTGA 62
QY 62 EKRAVTSPTMRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRIWTQ 121
DB 63 EKAAVTDVPTMRRGVSALAEASTAQVNTGTTDYSMSYSGMISGNVFPSEFVRVWTE 122
QY 122 YFDROYTASRAVAREVLRTATGTEPDGGVEAFDCBELLRFRYPQVPEHRSAAEQPL 178
DB 123 YFDKLYAAQETARLVLTASGGYDAEIVGSLDELDDADPVLRLRYPPEVPEHRSAAEPR 182
QY 179 RNAPHYDLSMVTLIQOTPCANGFVSLOAEVGAFTDLPYRDPDAVLVFCGAIATLVTGGQV 238
DB 183 RNAPHYDLSMVTLIQOTPCANGFVSLOAEVGAFTDLPYRDPDAVLVFCGAIATLVTGGQV 242
QY 239 KAPRHVAAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVDSLDGETATFQDWI 298
DB 243 PAPRHVRSFGAGMREGSDRTSSVFFLRPTTDFSFVAKARSYGLAVLDMETATFGDWI 302
QY 299 GGNVNI 305
DB 303 GTNYVTM 309
RESULT 5
EXPA_CEPAC STANDARD; PRT; 332 AA.
AC P11935;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cephalosporin biosynthesis expandase/hydroxylase [Includes:
DE Deacetoxycephalosporin C synthetase (EC 1.14.20.1) (DACS)
DE (Expandase); Deacetoxycephalosporin C hydroxylase (EC 1.14.11.-)
DE (Deacetylcephalosporin C synthetase) (DACS) (Beta-lactam
DE hydroxylase)].
GN CEFF.
OS Cephalosporium acremonium (Acremonium chrysogenum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Mitosporic Hypocreaceae;
OC Acremonium.
OX NCBI_TaxID=5044;
RN [1]
RP SEQUENCE FROM N.A.
RA Samson S.M., Dotzlauf J.E., Slisz M.L., Becker G.W., van Frank R.M.,
RA Vial L.E., Yeh W.K., Miller J.R., Queener S.W., Ingolia T.D.;
RT "Cloning and expression of the fungal expandase/hydroxylase gene
RT involved in cephalosporin biosynthesis.";
RL Biotechnology 5:1207-1214(1987).
CC
CC -1- FUNCTION: DACS catalyzes the step from penicillin N to deacetoxy-
CC cephalosporin C, which is used as a substrate by DACS to form
CC deacetylcephalosporin C.
CC -1- CATALYTIC ACTIVITY: Penicillin N + 2-oxoglutarate + O(2) =
CC deacetoxycephalosporin C + succinate + CO(2) + H(2)O.
CC -1- COFACTOR: Iron and ascorbate.
CC -1- PATHWAY: Cephalosporin biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES.
CC
CC PIR; A29711; A29711.
DR InterPro; IPR005123; 2OG-Fell_Oxy.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; 2OG-Fell_Oxy; 1.

```



```

RT complementation of the cephamycin pathway in Streptomyces
RL clavuligerus."
CC -1- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms
CC from delta-L-(alpha-aminoacidipyl)-L-cysteinyl-D-valine (ACV) to
CC form the azetidinone and thiazolidine rings of isopenicillin.
CC -1- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-
CC cysteinyl-D-valine + O(2) = isopenicillin N + 2 H(2)O.
CC -1- COFACTOR: Iron and ascorbate.
CC -1- PATHWAY: Biosynthesis of penicillin and cephalosporin.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X54609; CAA38431.1; -.
CC PIR; A61155; A61155.
CC HSP; P05326; 1BK0.
CC InterPro; IPR005123; 2OG-FeII_Oxy.
CC InterPro; IPR002283; IPN_synth.
CC Pfam; PF031171; 2OG-FeII_Oxy; 1.
CC PRINTS; PR00682; IPNSYNTHASE.
CC PROSITE; PS00185; IPNS_1; 1.
CC PROSITE; PS00186; IPNS_2; 1.
CC Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C.
FT METAL 212 212 IRON (BY SIMILARITY).
FT METAL 214 214 IRON (BY SIMILARITY).
FT METAL 268 268 IRON (BY SIMILARITY).
FT SEQUENCE 329 AA; 37368 MW; 0CD96C87C7CF5A7EB CRC64;
CC
CC Query Match 10.8%; Score 176; DB 1; Length 329;
CC Best Local Similarity 24.5%; Pred.No.9.7e-09;
CC Matches 73; Conservative 39; Mismatches 126; Indels 60; Gaps 12;
CC
Qy 5 VPFFSLAELOGLHOD-----EPRCLRDKGLFYLDGTLDTLKSADLVDPFEHG 58
Db 10 VPTIDISGGDADDKRVAQEINKACRSFGFYASHHG1--DVQLLDKVVNVEHRTM 66
Qy 59 SEAEK-----RAVTSVPVTRRGFTGLESESTAQITNTGSYSDYSCYSGMTADNLPF-- 111
Db 67 TDEKYDLAINAYNKNPNRTNGY-----YNAVKGKAVESWCYLNPSFSEDPHQI 117
Qy 112 -SG---DFGRW-----TQYFPRQYTSARAVAREVLRTATGTPDGGVEA--- 151
Db 118 RSGTPMEHGINWPDEKRHRQFRPCEQYRDVPSLSKVLKRGFPALALG-KPEDFFDASLS 176
Qy 152 ---FLDCEPILRFYFQVPEHRSABE-QPLRMAPHYDLSWVTLIQOTPCANGFVSLOAE 207
Db 177 LADTSLAVTLIHVPLEYEDYPVKTGPDGDKLSFEDHLDVSMITVLFTQEVQN-----LQVE 232
Qy 208 VGGFTDLPYRPDAVLVFCGAIALTVLGGQVKAPRRHVAAPRRDQIAGSSRTSSVPFL 265
Db 233 TADGWQDLPTSGENFLVNCGYTMGYLTDVFPAPNHRV-----KFINAERLSLPFL 284
CC
RESULT 8
ID -IPNS_STRCL STANDARD; PRT; 329 AA.
AC P10621;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Isopenicillin N synthetase (EC 1.2.1.3.1) (IPNS) (Isopenicillin N
DE synthase).
GN PCBC.
OS Streptomyces clavuligerus.

```

```

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1901;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 27064 / DSM 738 / NRRL 3585;
RC MEDLINE=98212175; PubMed=3130293;
RA Leski B.K., Aharonowitz Y., Mevarech M., Wolfe S., Vining L.C.,
RA Westlake D.W.S., Jensen S.E.;
RA "Cloning and nucleotide sequence determination of the isopenicillin N
RT synthase gene from Streptomyces clavuligerus.";
RL Gene 62:187-196(1988)
CC -!- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms
CC from delta-L-(alpha-aminoadipyl)-L-cysteiny-D-valine (ACV) to
CC form the azetidinone and thiazolidine rings of isopenicillin.
CC -!- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-
CC cysteinyl-D-valine + O(2) = isopenicillin N + 2 H(2)O.
CC -!- COFACTOR: Iron and ascorbate.
CC -!- PATHWAY: Biosynthesis of penicillin and cephalosporin.
CC -!- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M19421; AAA26770.1; -.
DR EMBL; A01132; CAA00131.1; -.
DR PIR; A29894; A29894.
DR HSPB; P05326; 1BLZ.
DR InterPro; IPR005123; ZOG-Felll Oxy.
DR InterPro; IPR002283; IPN_synth.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; ZOG-Felll Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
DR PROSITE; PS00185; IPNS_1; 1.
DR PROSITE; PS00186; IPNS_2; 1.
KW Antibiotic biosynthesias; Oxidoreductase; Iron; Vitamin C.
FT METAL 212 212 IRON (BY SIMILARITY).
FT METAL 214 214 IRON (BY SIMILARITY).
FT METAL 268 268 IRON (BY SIMILARITY).
SQ SEQUENCE 329 AA; 36958 MW; 71AA1CCE9514761C CRC64;
Query Match 10.08; Score 164; DB 1; Length 329;
Best Local Similarity 23.8; Pred.No.1.2e-07;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;
QY 5 VPTFSLAEI-----QQGLHQDEFRCLRDKGLFYLTDCGLTDTEKSAKDVLIDFFEHG 58
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 10 VPTIDISPLFGTDAAKRVAEIHNGCRSGFFYATNHGV---DVQQLQDVNFEGHGM 66
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 59 SEAEK-----RAVTSPVPTMRRGF----TCLESESTAQITNTGSYSYSCMYSGMT---A 106
DB ::|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 67 TDQEKHLATHAYNPDPNPHVRNGYKAVPGKVAESFCYLNPDPGEDHPW-IAAGTPME 125
DB ::|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 107 DNLPSPGD----FGRITWTQYFDROYTASRAVAEVL RATGETPDGGVEA-----FLDCE 156
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 126 VNLPWDPEERHPRFPFCPEGYYRQMLKLSTVLMRGALALG-RPEHFHFDAAEAQEQSLSVV 184
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 157 PLLRFRYFPQVBEHSABE-QPLRMAPHYDLNVTLIQPPPCANGFVSLOAEVGGAFTDL 215
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 185 SLRIPLYLEYPPVKTGPDGQLLSFEDHLDVSMITLVLFQTQVN----LQVETVDGHRDI 240
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 216 PYRPDAVLVFCGAIALTVTGQGVKAPRHVAAPRRDQIAGSSRSSTSSVFLLRPNAFTFSV 275
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 241 PTSENDFLVNGCYMAHTNDYPAPNHRV-----KFVVAERLSLPFFLNGGHEAVIE- 293
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 276 PLARECGDFVSLDGETAFTQDWI 298
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

```

```

Db 294 PFVPE-GASEVRNEALSVDYL 315
RESULT 9
IPNS_FLASS STANDARD; PRT; 326 AA.
AC P16020;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Isopenicillin N synthetase (EC 1.21.3.1) (IPNS) (Isopenicillin N
DE synthase).
DE PCBC.
OS Flavobacterium sp. (strain SC 12,154).
OC Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales;
OC Flavobacteriaceae; Flavobacterium.
OX NCBI_TaxID=241;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174998; PubMed=2308852;
RA Shiffman D., Cohen G., Aharonowitz Y., von Dohren H., Kleinkauf H.,
RA Mevarech M.;
RT "Nucleotide sequence of the isopenicillin N synthase gene (pcbC) of
RT the Gram-negative Flavobacterium sp. SC 12,154."
RL Nucleic Acids Res. 18:660-660(1990).
CC -I- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms
CC from delta-L-(alpha-aminoacyl)-L-cysteiny-D-valine (ACV) to
CC form the acetidinone and thiazolidine rings of isopenicillin.
CC -I- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-
CC cysteinyl-D-valine + O(2) = isopenicillin N + 2 H(2)O.
CC -I- COFACTOR: Iron and ascorbate.
CC -I- PATHWAY: Biosynthesis of penicillin and cephalosporin.
CC -I- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X17355; CAA35233.1; -.
CC DR HSP; P05326; 1BK0.
CC DR InterPro; IPR005123; 2OG-FeII Oxy.
CC DR InterPro; IPR002283; IPN synth.
CC DR Pfam; PF03171; 2OG-FeII Oxy; I.
CC DR PRINTS; PR00682; IPNSYNTHASE.
CC DR PROSITE; PS00185; IPNS_1; 1.
CC DR PROSITE; PS00186; IPNS_2; 1.
CC DR METAL 209 209 IRON (BY SIMILARITY).
CC DR METAL 211 211 IRON (BY SIMILARITY).
CC DR METAL 265 265 IRON (BY SIMILARITY).
CC DR SEQUENCE 326 AA; 36465 MW; 2568274316395837 CRC64;

Query Match 9.9%; Score 162; DB 1; Length 326;
Best Local Similarity 22.9%; Pred. No. 1.9e-07;
Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;

QY 24 RCLRDGLFYLTDGLTDELKSAKDLVDFPEHSGEAEK-----RAVTSPTVPTMRGFT 78
Db 32 RACRSGGFFFAAHNGV---DLAALQKFTTDMHMAAEKWEALIRAYNPANPRNNGY- 87
79 GLESESTAQITNTGYSYDSYCMYSGMTAD-----NLFPS-----GDFGRW 119
Db 88 -----YNAVEGKANESFCYLNPSFDADHATIKAGLPSHEVNIWPEARHPGMRFY 139
QY 120 TQVFDQRYTASRAVAEVLNATGTEPDGGVEAF-----LDCEPLLRFYFQVPE 169
Db 140 EAYFSDVFDVAAILRGAIFAALGRE-----ESFPERHFSMDDTLSAVSLIRYPFLENYP- 193

```

```

QY 170 HRSABEQPLRMAP-----HYDLSMTVLIQQTFCANGFVSLQAEVGGAFDTLFPRED 220
Db 194 -----PLKGPDEKLSFEHHQDVSLITLVLYQTAIPN-----LQVETAEGLDIPVSD 242
QY 221 AVLVFCAIATLVGGQVKAPRHVAAARRDQIAGSSRTSSVFF 264
Db 243 HFLVNGCTYMAHITNGYYPAPVHRV-----KYINAERLSIDPF 280

RESULT 10
IPNS_LYSLA STANDARD; PRT; 326 AA.
AC Q48739;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Isopenicillin N synthetase (EC 1.21.3.1) (IPNS) (Isopenicillin N
DE synthase).
DE PCBC.
OS Lysobacter lactamgenus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Lysobacter.
OX NCBI_TaxID=39596;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=YK90.
RX MEDLINE=96287475; PubMed=8703429;
RA Kimura H., Izawa M., Sumino Y.;
RT "Molecular analysis of the gene cluster involved in cephalosporin
RT biosynthesis from Lysobacter lactamgenus YK90."
RL Appl. Microbiol. Biotechnol. 44:589-596(1996).
CC -I- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms
CC from delta-L-(alpha-aminoacyl)-L-cysteiny-D-valine (ACV) to
CC form the acetidinone and thiazolidine rings of isopenicillin.
CC -I- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-
CC cysteinyl-D-valine + O(2) = isopenicillin N + 2 H(2)O.
CC -I- COFACTOR: Iron and ascorbate.
CC -I- PATHWAY: Biosynthesis of penicillin and cephalosporin.
CC -I- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X56660; CAA39983.1; -.
CC DR FIR; S54099; S54099.
CC DR HSP; P05326; 1BK0.
CC DR InterPro; IPR005123; 2OG-FeII Oxy.
CC DR InterPro; IPR002283; IPN synth.
CC DR InterPro; IPR002057; Isopen N synth.
CC DR Pfam; PF03171; 2OG-FeII Oxy; I.
CC DR PRINTS; PR00682; IPNSYNTHASE.
CC DR PROSITE; PS00185; IPNS_1; 1.
CC DR PROSITE; PS00186; IPNS_2; 1.
CC DR METAL 209 209 IRON (BY SIMILARITY).
CC DR METAL 211 211 IRON (BY SIMILARITY).
CC DR METAL 265 265 IRON (BY SIMILARITY).
CC DR SEQUENCE 326 AA; 36564 MW; 2573167173F8447 CRC64;

Query Match 9.9%; Score 162; DB 1; Length 326;
Best Local Similarity 22.9%; Pred. No. 1.9e-07;
Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;

QY 24 RCLRDGLFYLTDGLTDELKSAKDLVDFPEHSGEAEK-----RAVTSPTVPTMRGFT 78
Db 32 RACRSGGFFFAAHNGV---DLAALQKFTTDMHMAAEKWEALIRAYNPANPRNNGY- 87

```

QY 79 GLESESTAIQITNGSYSDYSCYMGATAD-----NLFFS-----GDFGRW 119
 Db 88 -----YMAVEGKANESFCYLNPSFDADHATIKAGLPSHEVNIWPDARHPGMRFPY 139
 QY 120 TQYDFRQYTRASAVAREVLBATGTEDPGGVEAF-----LDCEPLRFRYPQVPE 169
 Db 140 EAYFSDVDFVAAILRGAFAIALGRE-----ESFFERHFSMDTSLSAVSLIRYPFLENYP- 193
 QY 170 HRSAGEOPLRMAP-----HYDLSMTVLIQOTPCANGFVSLQAEVGGAFDLPVRPD 220
 Db 194 -----PLKLPDGEKLSFEHQDVSILVILVQTAPN-----LQVETAGYLDIPVSDE 242
 QY 221 AVLVFCAIATLVGGQVKAAPRHVAAAPRRDQIAGSSRTSSVFF 264
 Db 243 HFLVNCQTYMAHITNGYYPAPVHRV-----KVINAERLSIPFF 280

RESULT 11
 IPNS STRLP STANDARD; PRT; 333 AA.
 AC P12438;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Isopenicillin N synthetase (EC 1.21.3.1) (IPNS) (Isopenicillin N synthetase)
 GN PCBC.
 OS Streptomyces lipmanii (Streptomyces alboniger)
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=132472;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88314868; PubMed=3045077;
 RA Weigel B.J., Burgett S.G., Chen V.J., Skatrud P.L., Froluk C.A., Queener S.W., Ingolia T.D.;
 RT "Cloning and expression in Escherichia coli of isopenicillin N synthetase genes from Streptomyces lipmanii and Aspergillus nidulans";
 RT J. Bacteriol. 170:3817-3826(1988).
 RL J. Bacteriol. 170:3817-3826(1988).
 CC -1- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms from delta-L-(alpha-aminoadipyl)-L-cysteiny-D-valine (ACV) to form the azetidinone and thiazolidine rings of isopenicillin.
 CC -1- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-cysteinyl-D-valine + O(2) = isopenicillin N + 2 H(2)O.
 CC -1- COFACTOR: Iron and ascorbate.
 CC -1- PATHWAY: Biosynthesis of penicillin and cephalosporin.
 CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF OXIDOREDUCTASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M22081; AAA26771.1; .
 DR HSP; P05326; 1BK0.
 DR InterPro; IPR005123; 2OG-Fell_Oxy.
 DR InterPro; IPR002283; IPN synth.
 DR InterPro; IPR002057; isopen_N synth.
 DR Pfam; PF03171; 2OG-Fell_Oxy; 1.
 DR PRINTS; PR00682; IPNSYNTHASE.
 DR PROSITE; PS00185; IPNS 1; 1.
 DR PROSITE; PS00186; IPNS 2; 1.
 KW Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C.
 FT METAL 216 216 IRON (BY SIMILARITY).
 FT METAL 218 218 IRON (BY SIMILARITY).
 FT METAL 272 272 IRON (BY SIMILARITY).
 SQ SEQUENCE 333 AA; 38082 MW; D55385664EA2CA26 CRC64;

Query Match 8.7%; Score 143; DB 1; Length 333;
 Best Local Similarity 23.2%; Pred. NO. 1.1e-05;
 Matches 70; Conservative 37; Mismatches 131; Indels 64; Gaps 13;

QY 5 VPTFSLAELOOGLHQP-----EPRCLRDKGLFVLTDCGLTDTLKSADKOLVIDPFEH 57
 Db 10 VPTIDISPL-FGTDPDAKAVARQINEACRGSGFFYASHHGI---DVRRLQDVVNEFHRT 65
 QY 58 GSEAEK-----RAVTSPTVMRGFTGLESESTAIQITNGSYSDYSCY---SMGTADNL 109
 Db 66 MDOEKHDLAIHAYNENSHVNGY-----YWARFGRKTVESWCYLNPSFGEDHPM 116
 QY 110 FPSG-----DFGRITWQYDFRQYTRASRA-VAREVLRAVGT-----EPDG 147
 Db 117 IKAGTPMHEVNVMPDEERHPDERSFGQYREVRFLSKVLLRGLFALALGKPEPFENEV 176
 QY 148 GVEAFDCEPL-LRFRYPQVPE---HRSAGEOPLRMAPHYDLSMTVLIQOTPCANGFVS 203
 Db 177 TEEDTLSCRSMLIRYPYLDYPPEAAIKTGPDTGLSFEDHLDVSMITVLVFTQEVQN--- 232
 QY 204 LQAEVGGAFDLPYRDAVLVFCGAIATLVGGQVKAAPRHVAAAPRRDQIAGSSRTSSVF 263
 Db 233 LQVETVGGQSUFTSGENFLINCCTYGLVLTNDYFPAPNHRV-----KYVNAERLSLPP 286
 QY 264 FL 265
 Db 287 FL 288

RESULT 12
 FLS_PETH STANDARD; PRT; 348 AA.
 ID AC Q07512;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Flavanol synthase (EC 1.14.11.-) (FLS).
 GN FL.
 OS Petunia hybrida (Petunia).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Petunia.
 OX NCBI_TaxID=4102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Old Glory Blue; TISSUE=Petal;
 RX MEDLINE=94108485; PubMed=7904213;
 RA Holton T.A., Brugliera F., Tanaka Y.;
 RT "Cloning and expression of flavonol synthase from Petunia hybrida."; Plant J. 4:1003-1010(1993).
 RL Plant J. 4:1003-1010(1993).
 CC -1- FUNCTION: CATALYZES THE FORMATION OF FLAVONOLS FROM DIHYDROFLAVONOLS. IT CAN ACT ON DIHYDROKAEEMPEROL TO PRODUCE KAEEMFEROL, ON DIHYDROQUERCETIN TO PRODUCE QUERCETIN AND ON DIHYDROMYRICETIN TO PRODUCE MYRICETIN.
 CC -1- COFACTOR: IRON AND ASCORBATE.
 CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR FLAVONOLS WHICH ARE FLAVONOIDS. FLAVONOLS ARE IMPORTANT CO-PIGMENTS IN FLOWER OR FRUIT COLOR AND ARE ALSO ESSENTIAL FOR POLLEN TUBE GROWTH.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGHEST LEVEL DURING THE FIRST STAGE OF FLOWER DEVELOPMENT.
 CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF OXIDOREDUCTASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----


```
DR EMBL; Z22543; CAA80264.1; -.
DR PUR; S33510;
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
KW Flavonoid biosynthesis; Oxidoreductase; Dioxxygenase; Iron; Vitamin C.
SQ SEQUENCE 348 AA; 39427 MW; B39E1E4381DE6379 CRC64;

Query Match      8.6%; Score 140; DB 1; Length 348;
Best Local Similarity 22.3%; Pred. No. 2.2e-05;
Matches 59; Conservative 43; Mismatches 114; Indels 48; Gaps 9;

QY 27 RDGLFLVLTDCGLTDLTKSAKDLVDFRHSAGEAKRAVT-SPVPTMRGRF-TGLESES 84
Db 77 KENGIFQLINHGIPDENAIADLQVKGEFFHEVQEEKELIAKTPGNDIEGYTSLQKEV 136
QY 85 TAQITNTGSDYSYCMYSGMTADNLF-----PSGDFGRW-----TOYFDRQYT 128
Db 137 EKG-----KGMVDHLFHKIWPSSAVNRYWPKNPPSYREANEVEYKRMRE 181
QY 129 ASRAVAREVLRAATGPDGVEAFDCE--PLIRFRYPQVQVPEHRSABEOPLRMAPHYDL 186
Db 182 VDRIRFKSLGLGLEGHEMIEAAGDEIVYLLKINYPPCPR---PDALGVAHTDM 237
QY 187 SMYTLTQQTCCANGFVSQAEGVGAFTDLPYRPDAVLVFCGATATLVTTGGQVKAPRHVA 246
Db 238 SYTIIIL-----VPEVQGLQVFKDGHVYDVXKYNALIVHGDQVEILSNGKYKSVYHRT 293
QY 247 APRDQIAGSSRTSSVFFLRPNAD 270
Db 294 VNK-----DKTRMSWPVLEPPSE 312

RESULT 13
IPNS_NOCLA
ID FLS CITUN STANDARD; PRT; 328 AA.
AC P27744;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Isopenicillin N synthetase (EC 1.21.3.1) (IPNS) (Isopenicillin N synthase).
GN PCBC.
OS Nocardia lactamurans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VAR LC 411;
RX MEDLINE=92065808; PubMed=1956290;
RA Coque J.J.R., Martin J.F., Calzada J.G., Liras P.;
RT "The cephamycin biosynthetic genes pcbAB, encoding a large multidomain peptide synthetase, and pcbC of Nocardia lactamurans are clustered together in an organization different from the same genes in Actinonidium chrysogenum and Penicillium chrysogenum."
RT in Actinonidium chrysogenum and Penicillium chrysogenum."
RL Mol. Microbiol. 5:1125-1133(1991).
CC -1- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms from delta-L-(alpha-aminoadipyl)-L-cysteiny-D-valine (ACV) to form the acetidinone and thiazolidine rings of isopenicillin.
CC -1- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-cysteiny-D-valine + O(2) -> isopenicillin N + 2 H(2)O.
CC -1- COFACTOR: Iron and ascorbate.
CC -1- PATHWAY: Biosynthesis of penicillin and cephalosporin.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF OXIDOREDUCTASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:34:58 ; Search time 41 Seconds
(without alignments)
1957.423 Million cell updates/sec

Title: US-09-582-486-1
Perfect score: 1637
Sequence: 1 MDTTFTFSLAEQLQGLHOD.....ATFQDWIGGNVYNIIRTSKA 311

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Result No.	Score	Match	Query Length	DB ID	Description
1	1369	83.6	311	2 Q93FD4	Q93fd4 streptomyc
2	937	57.2	319	2 Q48740	Q48740 lysobacter
3	917.5	56.0	332	3 Q9P4T5	Q9p4t5 cephalospor
4	852	52.0	313	2 Q48741	Q48741 lysobacter
5	168	10.3	366	16 Q987V6	Q987v6 rhizobium 1
6	161	9.8	329	2 Q9ANU0	Q9anu0 streptomyc
7	159	9.7	282	2 Q9FAC2	Q9fac2 streptomyc
8	156.5	9.6	403	5 Q9VCE2	Q9vc22 drosophila
9	156.5	9.6	403	5 Q8MZG6	Q8mzg6 drosophila
10	156	9.5	262	2 Q9FAC0	Q9fac0 streptomyc
11	154.5	9.4	366	10 Q942Y0	Q942y0 oryza sativ
12	154	9.4	262	2 Q9FAC1	Q9fac1 streptomyc
13	154	9.4	262	2 Q9FAB6	Q9fab6 streptomyc
14	152	9.3	262	2 Q9FAC3	Q9fac3 streptomyc
15	152	9.3	379	10 Q8VKL2	Q8vxl2 fagus sylv
16	151	9.2	262	2 Q9FAB8	Q9fab8 streptomyc

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

17	151	9.2	262	2 Q9FAB9	Q9fab9 streptomyc
18	151	9.2	331	2 Q9LCZ2	Q9lc22 streptomyc
19	148.5	9.1	389	10 Q8RVF5	Q8rvf5 oryza sativ
20	146.5	8.9	362	10 Q80851	Q80851 arabidopsis
21	145.5	8.9	348	10 Q8SB89	Q8sb89 oryza sativ
22	144.5	8.8	376	10 Q9C955	Q9c955 arabidopsis
23	142.5	8.7	358	10 Q80850	Q80850 arabidopsis
24	140	8.6	262	2 Q9FAB7	Q9fab7 streptomyc
25	139.5	8.5	368	10 Q94LP7	Q94lp7 oryza sativ
26	136.5	8.3	380	10 Q39112	Q39112 arabidopsis
27	135.5	8.3	221	2 Q9F5K9	Q9f5k9 streptomyc
28	135.5	8.3	352	10 Q9LWJ3	Q9lwj3 oryza sativ
29	135	8.2	364	10 Q9LIF4	Q9lif4 arabidopsis
30	133	8.1	313	10 Q8H619	Q8h619 oryza sativ
31	131.5	8.0	355	16 Q8ZG11	Q8zgil versinia pe
32	131	8.0	378	10 Q39111	Q39111 arabidopsis
33	130	7.9	309	16 Q8D029	Q8d029 versinia pe
34	128.5	7.8	342	16 Q98NJ5	Q98nj5 rhizobium 1
35	128	7.8	346	10 Q8LP22	Q8lp22 nierenbergi
36	125	7.6	350	10 Q8H620	Q8h620 oryza sativ
37	125	7.6	382	10 Q04280	Q04280 phaseolus v
38	125	7.6	387	10 Q9LWJ4	Q9lwj4 oryza sativ
39	124.5	7.6	354	10 Q8SB87	Q8sb87 oryza sativ
40	124.5	7.6	370	10 Q04282	Q04282 phaseolus v
41	124	7.6	389	10 Q8L6W4	Q8l6w4 beta vulgar
42	124	7.6	392	10 Q9FS09	Q9fs09 malus domes
43	123.5	7.5	349	10 Q9LDV9	Q9ldv9 oryza sativ
44	123	7.5	339	10 Q40063	Q40063 hordeum vul
45	122.5	7.5	386	10 Q39541	Q39541 cucurbita m

ALIGNMENTS

RESULT 1

Q93FD4	PRELIMINARY;	PRT;	311 AA.
ID	Q93FD4		
AC	Q93FD4;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)		
DE	Deacetoxycephalosporin C synthase.		
OS	Streptomyces jumoniensis.		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
OC	Streptomycineae; Streptomycetaceae; Streptomyces.		
OX	NCBI_TaxID=1945;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 29864;		
RA	Sim T.S., Sim J.;		
RT	"Cloning and purification of Streptomyces jumoniensis		
RT	deacetoxycephalosporin C synthase.";		
RL	Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; AF317908; AAL09460.1; -		
DR	InterPro; IPR005123; 2OG-Fell_Oxy.		
DR	InterPro; IPR002057; Isopen_N_synth.		
DR	Pfam; PF03171; 2OG-Fell_Oxy; 1.		
DR	PROSITE; PS00186; IPNS_2; 1.		
KW	Porin.		
SQ	SEQUENCE 311 AA; 34242 MW; B821ED796A4F5C1A CRC64;		

Query Match 83.6%; Score 1369; DB 2; Length 311;
Best Local Similarity 81.4%; Pred. No. 1.4e-119;
Matches 253; Conservative 28; Mismatches 30; Indels 0; Gaps 0;

QY	1	MDTTFTFSLAEQLQGLHODFRRCRLDRKGLFYLTDCGLTDTLTKSAKDLVDFHHGSE	60
Db	1	MDTTFTFSLAEQLQGLHODEFRSCLAEGLFYLTDSLGLSDADQKSAKDAIDFFEHGTE	60
QY	61	AEKRAVTSVPVMTNRGFTGLESTAIQTNTGSYDSYSCYNSGTADNLPFGSGDFGRWT	120
Db	61	EKRAATSTIPTIRRGFTGLESTAIQTNTGSYDSYSCYNSGLADNVFPFGDFERVWT	120


```

Db      117  IIRSGTGMHEVINWPKDEKHERFRPFCEQYYRDMFQSLKLMRGFALALG-KPEDFFDANL 175
QY      152  ----FLDCPLLRFRFPQVBEHRSABE-QPLEMAHYDLSWVTLQQTPCANGFVSLQA 206
Db      176  PEDDTLSAVSLIRYPHKLKAYPPVKTGPDGTGKUSFEDHLDVSVITLVLFQTEVQN---LQV 231
QY      207  EVGGAFTDLPYRDVAVLFCGAIATLVGTQVKAPRRHVAAPRRDQIAGSSRTSSVFFL 265
Db      232  ETVNGHQDLPTSGDDFLVNGCTYMGVLTNDYPPAPNHRV-----KFINARLSLPPFL 284

RESULT 7
Q9FAC2  PRELIMINARY; PRT; 262 AA.
AC      Q9FAC2;
DT      01-MAR-2001 (TREMBlrel. 16, Created)
DT      01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE      Isopenicillin N synthase (IPNS) (Fragment).
GN      PCBC.
OS      Streptomyces panayensis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Streptomycineae; Streptomycetaceae; Streptomycetes.
OX      NCBI_TaxID=132470;
RN      [1]
RC      STRAIN=JCM 5042;
RA      Palaniappan N., Seki T.;
RT      "Phylogenetic relationship between the Isopenicillin N synthase (IPNS)
RL      and 16S rDNA genes.";
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AB045851; BAB13299.1; -
DR      HSSP; P05326; 1BK0.
DR      InterPro; IPR005123; 2OG-Fell Oxy.
DR      InterPro; IPR002283; IPN synth.
DR      InterPro; IPR002057; Isopen N synth.
DR      Pfam; PF031171; 2OG-Fell Oxy; 1.
DR      PRINTS; PR00682; IPNSYNTHASE.
DR      PROSITE; PS00185; IPNS 1; 1.
DR      PROSITE; PS00186; IPNS 2; 1.
FT      NON TER 1
FT      NON TER 262
SQ      SEQUENCE 262 AA; 30074 MW; D9A5D0BAB942D776 CRC64;

Query Match 9.7%; Score 159; DB 2; Length 262;
Best Local Similarity 25.4%; Pred. No. 1.4e-06;
Matches 63; Conservative 38; Mismatches 115; Indels 32; Gaps 10;

QY      21  EFRCLRDKGLFYLTDCGLTDTBLKSAKDLIVDFEHSSEAEK-----RAVTSVPVPTMR 75
Db      17  EINKACRGSGFFVASHHGI---DVQLLDVVNEFHRTMTDEEKYDLAINAYKNKDNHVN 73
QY      76  GF-----TGLESESTAOITNGSYSDYSCVSMGT---ADNLFPS-----GDFGRINTQYFD 124
Db      74  GYTMVAVKGAKEVSWCYLNPFSFSEDFMIRS-GTPMHEVINWPKDEKRAHFRPFCEQYR 132
QY      125  RQYTSARAVAREVLRTATGTEPDGGVEA-----FLDCEPLLRFRFPQVBEHRSABE-QP 177
Db      133  DMFQLSKALMRGFPALLG-KPEDFEAPLSQDFTLSAVSMIRPYLDEPPVKTGPDGTGK 191
QY      178  LRMAPHYDLSWVTLIQQTPCANGFVSLQAEVGAFTDLPYRDVAVLFCGAIATLVGTGQ 237
Db      192  LSFEDHLDVSMITVLVLFQTEVQN---LQVETVDGWDQLPTSGENFLVNGCTYMGVLTNDY 247
QY      238  VKAPRRHV 245
Db      248  FPAPNHRV 255

RESULT 8
Q9VCZ2  PRELIMINARY; PRT; 403 AA.
ID      Q9VCZ2
AC      Q9VCZ2;

```

```

DT      01-MAY-2000 (TREMBlrel. 13, Created)
DT      01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      CG5346 protein.
GN      CG5346.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RC      SEQUENCE FROM N.A.
RC      STRAIN=Berkelley;
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA      Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA      Abril J.F., Agbavani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA      Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA      Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA      Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA      Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA      Foster K.C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA      Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA      Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA      Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA      Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA      Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA      Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA      Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA      Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA      Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA      Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA      Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Ye J.,
RA      Williams S.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA      Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA      Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA      Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA      Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,
RA      Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA      Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA      McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA      Pacleb J., Paragias V., Park S., Patel S., Pfeiffer B.,
RA      Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA      Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,
RA      Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT      "Sequencing of Drosophila melanogaster genome."
RL      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN      [3]
RC      SEQUENCE FROM N.A.
RP      Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA      Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,

```

RA	Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA	Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA	Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA	Searle S.M.J., Smith E., Shu S., Smutniak P., Whitfield E.,
RA	Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT	"Annotation of Drosophila melanogaster genome.";
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN	[4]
RN	SEQUENCE FROM N.A.
RP	Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL	[5]
RP	SEQUENCE FROM N.A.
RA	FlyBase;
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AE003739; AAF56013.2; -.
DR	FlyBase; FBgn0038981; CG5346.
DR	InterPro; IPR005123; 2OG-Fell Oxy.
DR	Pfam; PF03171; 2OG-Fell Oxy; I.
DR	SEQUENCE 403 AA; 45431 MW; DE2CA70B6443A07C CRC64;
QY	Query Match 9.6%; Score 156.5; DB 5; Length 403;
Db	Best Local Similarity 20.2%; Pred. No. 4.3e-06;
Matches	72; Conservative 51; Mismatches 127; Indels 107; Gaps 12;
QY	3 TTUPTFSLAELOOGLHQ-----DEPRCLRDKGLPYLTDCGLTDELKSADLV 51
Db	: : : : : : : : : : : : : : :
18	SVVPIIDLAHC-GIEEPVKVSVVNRVGHQLKALSEKMALLYNHGISEDKLATWDHL 75
QY	52 IDPFEHSGEAKRAVTSPVPMTRRGFTGLESESTAQITN-----TGSYSDYSYCSMGT 105
Db	:: :: :: :: :: :: :: :
76	DDP-----VNLP-PDIRHYRADGKHGYVSRGQQQRFDGKSPELRHAFNI 123
QY	106 --ADNL-----FPSGDGRIMTYFDROQTASRAVAAREVLATGTEPDGGVEAFL----- 153
Db	: : : : : : : : :
124	LNAQNLPPEPLPG-----PADHISTLATDPKALASFILQALAVSLDIPHTFFLEKSHM 177
QY	154 -----DCEPLLRFYPQV-----PEHRSAERQPL 178
Db	: : : : : : : :
178	LSGDHDNMSSLRMLYPPIVDDEFGQNDVIKRCQYSYQRCLSNQOPFRPEHNPRDEDDL 237
QY	179 -----RMAPHYDLMSVTLTIQOTPCANGFVSQAQEVGAFTDLPY 217
Db	: : : : : : : :
238	NEVDGPNQOLFPEHKLNGAVIQCPEPHVDYGTFTLLSQD--SEGGLVELPGEKKNRVGH 295
QY	218 RPDVAVFCGAATLVTTGGQVKAPRHHVAAPRRDOIAGSSTTSVFVFLRNADTFPS 274
Db	: : : : : : : :
296	LPGSILVNCGEILNMTQGRVPALOHVIIPEOSTIRARGHSHTAFFCHPNITTIS 352

```

RESULT 9
Q8MZG6
ID Q8MZG6 PRELIMINARY; PRT; 403 AA.
AC Q8MZG6;
DT 01-OCT-2002 (TREMBlurel. 22, Created)
DT 01-OCT-2002 (TREMBlurel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlurel. 23, Last annotation update)
DE SD09582p.
GN CG5346.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
RN
RP SEQUENCE FROM N.A.
RA Stapleton M., Bröckstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Muncall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.
RL submitted (May-2002) to the EMBL/GenBank/DBJ databases.

```

DR	ENBL; AY102699; AAM27528.1; -;
DR	FlyBase; FBgn0038981; CG5346.
DR	InterPro; IPR005123; 2OG-FeII_Oxy.
DR	Pfam; PF03171; 2OG-FeII_Oxy; 1.
SQ	SEQUENCE 403 AA; 45461 MW; DE2CBC0FD443A07C CRC64;
 Query Match 9.6%; Score 156.5; DB 5; Length 403; Best Local Similarity 20.2%; Pred. No. 4.3e-06; Matches 72; Conservative 51; Mismatches 127; Indels 107; Gaps 12;	
QY	3 TTVPFTSLAELOOGLHQ-----DEFRCLRDKGLFYLTDCGLTDTLKSADLV 51 : 18 SVVPIIDLHC--GIEVPVKSVNVRGHQLKKALSEKGWALLVNHGISDEKLATANDHL 75 :
QY	52 IDPFGHGEAKKAIVTSVPVTMRGFTGLESESTAQTN-----TGSYSDYSVCYMGMT 105 : 76 DDF-----VNLP-PDIRQHVIYRADGDGHGVSRGQQORFDDGKSPELRHAFNIST 123 :
QY	106 --ADNL-----FPSGDFGRITWFDRQYTASRAVARAVLRATCTEPDGGVEAPL----- 153 : 124 LNAQNLPPEPLPG-----FADHISTLATDFKALASFILQALAVSLDIPIHTFFLEKHSHM 177 :
QY	154 -----DCEPLLRFYFPQV-----RMAPHYDLMSVTLTIQOTPCANGFSVLAQEVGAETDLPY 178 : 178 LSGDHNNWSLRMLYPPIVDDEPQGNDVIKRCQVSYQRCLSNQPDRFEHNPRDEDDL 237 :
QY	179 -----RMAPHYDLMSVTLTIQOTPCANGFSVLAQEVGAETDLPY 217 : 238 NEVDGPNGLQFEHKLNGNAVICTPPHYDVYGTFTLLSQD---SEGLEVLRLPGSEKNRVGH 295 :
QY	218 RPAVLVFCAIATLVTCGVKAPRHVAAPRDDIAGSSRTSSVFFLRPNADTFPS 274 : 296 LPGSILNVCGEILNIWTQGRYPALQHRVIIPEQETIRARGRHSTAFFCHPDNITTIS 352 :
 RESULT 10 Q9FAC0 PRELIMINARY; PRT; 262 AA. ID Q9FAC0 AC Q9FACO DT 01-MAR-2001 (TrEMBLrel. 16, Created) DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update) DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update) DE Isopenicillin N synthase (IPNS) (Fragment). GN PCBC. OS Streptomyces heteromorphus. OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; OC Streptomycineae; Streptomycetaceae; Streptomyces. OX NCBI_TaxID=132471; RN [1] RP SEQUENCE FROM N.A. RC STRAIN=JCM 5031; RA Palaniappan N., Seki T.; RT "Phylogenetic relationship between the Isopenicillin N synthase (IPNS) and 16 rDNA genes."; RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases. DR EMBL; AB045853; BAB13301.1; -; DR HSP; P05326; IBKO. DR InterPro; IPR005123; 2OG-FeII_Oxy. DR InterPro; IPR002283; IPN_synth. DR InterPro; IPR002057; Isopen_N_synth. DR Pfam; PF03171; 2OG-FeII_Oxy; 1. DR PRINTS; PS00682; IPNSYNTHASE. DR PROSITE; PS00185; IPNS_1; 1. DR PROSITE; PS00186; IPNS_2; 1. FT NON_TER 1 FT NON_TER 262 SQ SEQUENCE 262 AA; 30122 MW; 3AFED20F3DB41E69 CRC64;	
 Query Match 9.5%; Score 156; DB 2; Length 262; Best Local Similarity 25.0%; Pred. No. 2.6e-06; Matches 62; Conservative 38; Mismatches 116; Indels 32; Gaps 10	
QY	21 EFRRCLRDKGLFYLTDCGLTDTLKSADLVIDPFHGSEAEK-----RAVTSVPVTMMR 75

NCBI_TaxID=67274;
[1]
SEQUENCE FROM N.A.
STRAIN=JCM 4623;
Palaniappan N., Seki T.;
"Phylogenetic relationship between the Isopenicillin N synthase (IPNS) and 16S rDNA genes.";
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB045857; BAB13305.1; -.
DR HSSP; P05326; 1BK0.
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
DR PROSITE; PS00185; IPNS 1; 1.
DR PROSITE; PS00186; IPNS 2; 1.
FT NON_TER 1 1
SQ SEQUENCE 262 AA; 29991 MW; CFAA44B6D7D307B6 CRC64;
Query Match 9.4%; Score 154; DB 2; Length 262;
Best Local Similarity 24.3%; Pred. No. 4e-06;
Matches 62; Conservative 32; Mismatches 115; Indels 46; Gaps 9;
QY 21 EFRCLRDKGLFYLTDCGLTDTLKSADLVDPFEHGEAEK-----RAVTSVPVPTWRR 75
DB 17 QINKACRGSGFFVASHHGI---DVORLQDVVNEFHRTWDSQEKYDLAINAYNENPHVRN 73
QY 76 GFTGLESESTAQITNTGSDYSYMCY-----SMGT---ADNLFPSGD-----FG 116
DB 74 GY-----YNAVKGRKAVESFCYLNFLFNDEHPMKSGTPMHEVNIWPEDEDRHGFGR 124
QY 117 RIWTOYFDRQYTSRAVAREVLRTCT-----EPDGGVEAFDCEPLLRFRYFPQVPEHR 171
DB 125 TFCQYTRDMLQSTVIMRGIALALGAKENFFDPALAKADSLSSVSMIRIPLYEDYPPVK 184
QY 172 SAEQ-QPLRMAPHYDLSMTLLQOTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIA 230
DB 185 TGPDGTKLSFEDHLDVSMITLVLFQTEVQN---LQVETVDGQDLPTSDNFLVNCGYM 240
QY 231 TLVTGGQVQKAPRHV 245
DB 241 GHVTNDYFPAPNHRV 255
RESULT 14
Q9FAC3 PRELIMINARY; PRT; 262 AA.
ID Q9FAC3
AC Q9FAC3
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Isopenicillin N synthase (IPNS) (Fragment).
GN PCBC.
OS Streptomyces viridochromogenes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1938;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=JCM 5013;
RA Palaniappan N., Seki T.;
RT "Phylogenetic relationship between the Isopenicillin N synthase (IPNS) and 16S rDNA genes.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB045850; BAB13298.1; -.
DR HSSP; P05326; 1BK0.
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
FT NON_TER 1 1
SQ SEQUENCE 262 AA; 29991 MW; CFAA44B6D7D307B6 CRC64;
Query Match 9.3%; Score 152; DB 10; Length 379;
Best Local Similarity 24.3%; Pred. No. 1e-05;
Matches 67; Conservative 40; Mismatches 105; Indels 64; Gaps 12;
QY 30 GLFYLTDCGLTDTLKSADLVDPFEHGEAEKRAVTSVPVPTWRR-----GFTG- 79
DB 92 GFFLVVNHGVDKLIATAHAHQYIDYFELPMKAKORA-----QRKKVGEHCYASSFTGR 144
QY 80 -----LSESTAQITNTGSDYSYMCYMGATADNLPFGSGDFGRIMTQYFDRQYTA 129
DB 145 FSFKLPWKETLSFRSAQPDSSNIQDY-LCNTWG--EDFEP---FGKYQDYCDAMSTL 198
QY 130 SRVAREVLRTATGTEPDGGVEAFDCEPLLRFRYFP---QVPEHRSAEQPLRMAPHYDLS 187
NCBI_TaxID=67274;
[1]
SEQUENCE FROM N.A.
STRAIN=JCM 4623;
Palaniappan N., Seki T.;
"Phylogenetic relationship between the Isopenicillin N synthase (IPNS) and 16S rDNA genes.";
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB045857; BAB13305.1; -.
DR HSSP; P05326; 1BK0.
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
FT NON_TER 1 1
SQ SEQUENCE 262 AA; 30107 MW; D72BF9737CC8EAEBCRC64;
Query Match 9.3%; Score 152; DB 2; Length 262;
Best Local Similarity 25.0%; Pred. No. 6.2e-06;
Matches 62; Conservative 39; Mismatches 115; Indels 32; Gaps 10;
QY 21 EFRCLRDKGLFYLTDCGLTDTLKSADLVDPFEHGEAEK-----RAVTSVPVPTWRR 75
DB 17 EINKACRGSGFFVASHHGI---DVORLQDVVNEFHRTWDSQEKELAIHAYNQANPRVRN 73
QY 76 GF-----TGLESESTAQITNTGSDYSYMCYSGMT---ADNLFPSGD-----FGRIMTQYFD 124
DB 74 GYMAVKGKKAVERSMCYLNPAGFEDHPMIRS-GTFLHEVNIWPEDEKHERFRPFCEQYIR 132
QY 125 RQYTSRAVAREVLRTATGTEPDGGVEAFD---DCEPLLRFRYFPQVPEHRSAEQRP----- 177
DB 133 DMFGLSKTLMRGFALALG-KPEDFPDAHLPEADTILSAVSLIRYPRLEDYPPVKTPGDGDK 191
QY 178 LRMAHYDLSMTLLQOTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIALTLVTGGQ 237
DB 192 LSFQDHLVDVSMITLVLFQTEVQN---LQVETVDGQDLPTSGDHPFLVNCGTFMGVLTNDY 247
QY 238 VKAPRHV 245
DB 248 FPAPNHRV 255
RESULT 15
Q8VXL2 PRELIMINARY; PRT; 379 AA.
ID Q8VXL2
AC Q8VXL2
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Gibberellin 20-oxidase 1.
GN GA20OX1.
OS Fagus sylvatica (Beechnut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosidia I; Fagales; Fagaceae; Fagaceae; Fagus.
OX NCBI_TaxID=28930;
RN [1]
RN SEQUENCE FROM N.A.
RA Calvo A., Lorenzo O., Rodriguez D., Nicolas G.;
RT "Molecular cloning and transcript analysis of a functional GA20 oxidase related to the release from dormancy in Fagus sylvatica seeds.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ420192; CAD21846.1; -.
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
SQ SEQUENCE 379 AA; 43056 MW; E3E51A2FC94EFBB3 CRC64;
Query Match 9.3%; Score 152; DB 10; Length 379;
Best Local Similarity 24.3%; Pred. No. 1e-05;
Matches 67; Conservative 40; Mismatches 105; Indels 64; Gaps 12;
QY 30 GLFYLTDCGLTDTLKSADLVDPFEHGEAEKRAVTSVPVPTWRR-----GFTG- 79
DB 92 GFFLVVNHGVDKLIATAHAHQYIDYFELPMKAKORA-----QRKKVGEHCYASSFTGR 144
QY 80 -----LSESTAQITNTGSDYSYMCYMGATADNLPFGSGDFGRIMTQYFDRQYTA 129
DB 145 FSFKLPWKETLSFRSAQPDSSNIQDY-LCNTWG--EDFEP---FGKYQDYCDAMSTL 198
QY 130 SRVAREVLRTATGTEPDGGVEAFDCEPLLRFRYFP---QVPEHRSAEQPLRMAPHYDLS 187

NCBI_TaxID=67274;
[1]
SEQUENCE FROM N.A.
STRAIN=JCM 4623;
Palaniappan N., Seki T.;
"Phylogenetic relationship between the Isopenicillin N synthase (IPNS) and 16S rDNA genes.";
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB045857; BAB13305.1; -.
DR HSSP; P05326; 1BK0.
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
DR PROSITE; PS00185; IPNS 1; 1.
DR PROSITE; PS00186; IPNS 2; 1.
FT NON_TER 1 1
SQ SEQUENCE 262 AA; 29991 MW; CFAA44B6D7D307B6 CRC64;
Query Match 9.4%; Score 154; DB 2; Length 262;
Best Local Similarity 24.3%; Pred. No. 4e-06;
Matches 62; Conservative 32; Mismatches 115; Indels 46; Gaps 9;
QY 21 EFRCLRDKGLFYLTDCGLTDTLKSADLVDPFEHGEAEK-----RAVTSVPVPTWRR 75
DB 17 QINKACRGSGFFVASHHGI---DVORLQDVVNEFHRTWDSQEKYDLAINAYNENPHVRN 73
QY 76 GFTGLESESTAQITNTGSDYSYMCY-----SMGT---ADNLFPSSGD-----FG 116
DB 74 GY-----YNAVKRKAVESFCYLNLFNDEHPMKSGTPMHEVNIWPEDEDRHGFGR 124
QY 117 RIWTOYFDRQYTSRAVAREVLRTCT-----EPDGGVEAFDCEPLLRFRYFPQVPEHR 171
DB 125 TFCQYTRDMLQSTVIMRGIALALGAKENFFDPAKADSLSSVSMIRIPLYEDYPPVK 184
QY 172 SAEQ-QLRMAPHYDLSMTLLQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIA 230
DB 185 TGPDGTKLSFEDHLDVSMITLVLFQTEVQN---LQVETVDGQDLPTSDNFLVNCGYM 240
QY 231 TLVTGGQVKAHRHV 245
DB 241 GHVTNDYFPAPNHRV 255
RESULT 14
Q9FAC3 PRELIMINARY; PRT; 262 AA.
ID Q9FAC3
AC Q9FAC3
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Isopenicillin N synthase (IPNS) (Fragment).
GN PCBC.
OS Streptomyces viridochromogenes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1938;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=JCM 5013;
RA Palaniappan N., Seki T.;
RT "Phylogenetic relationship between the Isopenicillin N synthase (IPNS) and 16S rDNA genes.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB045850; BAB13298.1; -.
DR HSSP; P05326; 1BK0.
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
FT NON_TER 1 1
SQ SEQUENCE 262 AA; 29991 MW; CFAA44B6D7D307B6 CRC64;
Query Match 9.3%; Score 152; DB 10; Length 379;
Best Local Similarity 24.3%; Pred. No. 1e-05;
Matches 67; Conservative 40; Mismatches 105; Indels 64; Gaps 12;
QY 30 GLFYLTDCGLTDTLKSADLVDPFEHGEAEKRAVTSVPVPTWRR-----GFTG- 79
DB 92 GFFLVVNHGVDKLIATAHAHQYIDYFELPMKAKORA-----QRKKVGEHCYASSFTGR 144
QY 80 -----LSESTAQITNTGSDYSYMCYMGATADNLPFGSGDFGRIMTQYFDRQYTA 129
DB 145 FSFKLPWKETLSFRSAQPDSSNIQDY-LCNTWG--EDFEP---FGKYQDYCDAMSTL 198
QY 130 SRVAREVLRTATGTEPDGGVEAFDCEPLLRFRYFP---QVPEHRSAEQPLRMAPHYDLS 187
NCBI_TaxID=67274;
[1]
SEQUENCE FROM N.A.
STRAIN=JCM 4623;
Palaniappan N., Seki T.;
"Phylogenetic relationship between the Isopenicillin N synthase (IPNS) and 16S rDNA genes.";
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB045857; BAB13305.1; -.
DR HSSP; P05326; 1BK0.
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
FT NON_TER 1 1
SQ SEQUENCE 262 AA; 30107 MW; D72BF9737CC8EAEBCRC64;
Query Match 9.3%; Score 152; DB 2; Length 262;
Best Local Similarity 25.0%; Pred. No. 6.2e-06;
Matches 62; Conservative 39; Mismatches 115; Indels 32; Gaps 10;
QY 21 EFRCLRDKGLFYLTDCGLTDTLKSADLVDPFEHGEAEK-----RAVTSVPVPTWRR 75
DB 17 EINKACRGSGFFVASHHGI---DVORLQDVVNEFHRTWDSQEKELAIHAYNQANPRVRN 73
QY 76 GF-----TGLESESTAQITNTGSDYSYMCYSGMT---ADNLFPSSGD-----FGRIMTQYFD 124
DB 74 GYMAVKGKKAVERSMCYLNPAGFEDHPMIRS-GTFLHEVNIWPEDEKHERFRPFCEQYIR 132
QY 125 RQYTSRAVAREVLRTATGTEPDGGVEAFD---DCEPLLRFRYFPQVPEHRSAEQRP----- 177
DB 133 DMFGLSKTLMRGFALALG-KPEDFPDAHLPEADTILSAVSLIRYPRLEDPVPVKTPGDGDK 191
QY 178 LRMAPHYDLSMTLLQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIALTLVTGGQ 237
DB 192 LSFQDHLVDVSMITLVLFQTEVQN---LQVETVDGQDLPTSGDHPFLVNCGTFMGVLTNDY 247
QY 238 VKAPRHHV 245
DB 248 FPAPNHRV 255
RESULT 15
Q8VXL2 PRELIMINARY; PRT; 379 AA.
ID Q8VXL2
AC Q8VXL2
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Gibberellin 20-oxidase 1.
GN GA20OX1.
OS Fagus sylvatica (Beechnut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosidia I; Fagales; Fagaceae; Fagus.
OX NCBI_TaxID=28930;
RN [1]
RN SEQUENCE FROM N.A.
RA Calvo A., Lorenzo O., Nicolas C., Rodriguez D., Nicolas G.;
RT "Molecular cloning and transcript analysis of a functional GA20 oxidase related to the release from dormancy in Fagus sylvatica seeds.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ420192; CAD21846.1; -.
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
SQ SEQUENCE 379 AA; 43056 MW; E3E51A2FC94EFBB3 CRC64;
Query Match 9.3%; Score 152; DB 10; Length 379;
Best Local Similarity 24.3%; Pred. No. 1e-05;
Matches 67; Conservative 40; Mismatches 105; Indels 64; Gaps 12;
QY 30 GLFYLTDCGLTDTLKSADLVDPFEHGEAEKRAVTSVPVPTWRR-----GFTG- 79
DB 92 GFFLVVNHGVDKLIATAHAHQYIDYFELPMKAKORA-----QRKKVGEHCYASSFTGR 144
QY 80 -----LSESTAQITNTGSDYSYMCYMGATADNLPFGSGDFGRIMTQYFDRQYTA 129
DB 145 FSFKLPWKETLSFRSAQPDSSNIQDY-LCNTWG--EDFEP---FGKYQDYCDAMSTL 198
QY 130 SRVAREVLRTATGTEPDGGVEAFDCEPLLRFRYFP---QVPEHRSAEQPLRMAPHYDLS 187

Db	199	SLGIMELLGMSLGVSOQHYREFEENESINRLNYPYCQKP-----DLTGTGPHCDPT	252
Qy	188	MVTLIQOTPCANGFVSLQAEVGG--AFTDLPYRP-----DAVLVFCGAIATLVTGGQVKA	240
Db	253	SLTILHQD-----QVGGLOVFEDEWRSITPNFNAFVVNIGDTFMALSNGRYKS	301
Qy	241	PRHHVAAPRRDQIAGSSRTSSVFRLRPNADFTFSVP	276
Db	302	CLHRAVVNSK-----TPRKSLAFFLCPCKNDKVVSPP	332

Search completed: February 3, 2004, 17:38:34
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:37:09 ; Search time 21 Seconds
(without alignments)
626.604 Million cell updates/sec

Title: US-09-582-486-1

Perfect score: 1637

Sequence: 1 MDTVTFTSLAELOQGLHOD.....ATFQDWIGNYNVRTSKA 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
 - 2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
 - 3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
 - 4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
 - 6: /cgn2_6/ptodata/1/iaa/backfile1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1624	99.2	311	3	US-09-413-231-18
2	1201	73.4	313	1	US-08-592-411-15
3	1186.5	72.5	314	1	US-08-592-411-17
4	965.5	59.0	310	5	PCT-US95-04801-6
5	164	10.0	329	3	US-09-413-231-5
6	164	10.0	329	3	US-09-413-231-9
7	162	9.9	335	2	US-08-379-556A-10
8	149	9.1	333	2	US-08-379-556A-8
9	143	8.7	333	3	US-09-413-231-10
10	140	8.6	347	2	US-08-379-556A-2
11	137	8.4	321	3	US-09-413-231-8
12	133.5	8.2	331	3	US-09-413-231-4
13	131	8.0	378	2	US-08-553-367A-6
14	131	8.0	378	3	US-09-295-306-6
15	131	8.0	378	4	US-09-734-719-6
16	130.5	8.0	343	4	US-09-454-034-8
17	122.5	7.5	331	3	US-09-413-231-1
18	122.5	7.5	331	3	US-09-413-231-2
19	122.5	7.5	331	3	US-09-413-231-3
20	122.5	7.5	386	2	US-08-553-367A-2
21	122.5	7.5	386	3	US-09-295-306-2
22	122.5	7.5	386	4	US-09-734-719-2
23	120.5	7.4	377	2	US-08-553-367A-4
24	120.5	7.4	377	3	US-09-295-306-4
25	120.5	7.4	377	4	US-09-734-719-4
26	117	7.1	344	4	US-09-252-991A-28503
27	109	6.7	338	3	US-09-413-231-7

28	108.5	6.6	390	4	US-09-252-991A-20513	Sequence 20513, A
29	106	6.5	241	2	US-08-379-556A-4	Sequence 4, Appli
30	106	6.5	338	3	US-09-413-231-6	Sequence 6, Appli
31	98	6.0	377	5	PCT-US95-10403-4	Sequence 4, Appli
32	96	5.9	357	4	US-09-645-168-3	Sequence 3, Appli
33	95.5	5.8	412	3	US-09-027-064-4	Sequence 4, Appli
34	95.5	5.8	412	3	US-09-271-815-4	Sequence 4, Appli
35	94.5	5.8	112	4	US-09-454-034-10	Sequence 10, Appli
36	94	5.7	375	2	US-09-645-168-2	Sequence 2, Appli
37	93	5.7	358	4	US-08-823-986A-2	Sequence 2, Appli
38	93	5.7	358	3	US-08-784-385-2	Sequence 2, Appli
39	89.5	5.5	363	1	US-07-772-032-2	Sequence 2, Appli
40	89.5	5.5	363	3	US-09-157-077-9	Sequence 9, Appli
41	89.5	5.5	363	5	PCT-US92-08012-2	Sequence 2, Appli
42	89	5.4	377	2	US-08-823-986A-4	Sequence 4, Appli
43	89	5.4	377	3	US-08-784-385-4	Sequence 4, Appli
44	84.5	5.2	430	4	US-09-252-991A-22599	Sequence 22599, A
45	84	5.1	218	4	US-09-252-991A-26105	Sequence 26105, A

ALIGNMENTS

RESULT 1

US-09-413-231-18
; Sequence 18, Application US/09413231
; Patent No. 6284483

; GENERAL INFORMATION:

; APPLICANT: Dilley, David R

; APPLICANT: Kadyrzhanova, Dina K

; APPLICANT: Wang, Zhenyong

; APPLICANT: Warner, Toni M

; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE REFERENCE: Cephalosporins Under the Control of Bicarbonate

; CURRENT APPLICATION NUMBER: US/09/413,231

; CURRENT FILING DATE: 1999-10-06

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 18

; LENGTH: 311

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: modified DAOCS

; OTHER INFORMATION: from Streptomyces clavuligerus

; NAME/KEY: MUTAGEN

; LOCATION: (181)

; OTHER INFORMATION: Ala181 in native DAOCS modified to Arg

US-09-413-231-18

Query Match 99.2%; Score 1624; DB 3; Length 311;

Best Local Similarity 99.4%; Pred. No. 2.3e-176; Indels 0; Gaps 0;

Matches 309; Conservative 0; Mismatches 2;

Qy 1 MDTVTFTSLAELOQGLHODEFRRCCLRDGKLFYLTDCGLTDTLTKSAKOLVDFEHGSE 60

Db 1 MDTVTFTSLAELOQGLHODEFRRCCLRDGKLFYLTDCGLTDTLTKSAKOLVDFEHGSE 60

Qy 61 AEKRAVTSVPVPMRRGFTGLESESTAQINTGTSYSDYSCYSGMTADNLPSPGDPGRWT 120

Db 61 AEKRAVTSVPVPMRRGFTGLESESTAQINTGTSYSDYSCYSGMTADNLPSPGDPGRWT 120

Qy 121 QYFDROYTASRAVAREVLRTATGTPDGGVEAFDCEPLLRFRFYFPOVPEHRSABEQPLRM 180

Db 121 QYFDROYTASRAVAREVLRTATGTPDGGVEAFDCEPLLRFRFYFPOVPEHRSABEQPLRM 180

Qy 181 APHYDLSMTVLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVGGQVKA 240

Db 181 RPHYDLSMTVLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVGGQVKA 240

Qy 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGDFVSLDGTATFDQWIGG 300

Db 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFVPLARECGDFVSLDGTATFDQWIGG 300

```
Db 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLAARECGFDVSLDGETATFDWIGG 300
QY 301 NYNIRRTSKA 311
Db 301 NYNIRRTSKA 311

RESULT 2
US-08-592-411-15
; Sequence 15, Application US/08592411
; Patent No. 5726032
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Process for the Efficient Production of
; TITLE OF INVENTION: 7-ADCA via 2-(Carboxyethylthio)acetyl-7-ADCA and
; TITLE OF INVENTION: 3-(Carboxymethylthio)propionyl-7-ADCA
; NUMBER OF SEQUENCES: 17
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,411
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-592-411-15

Query Match 73.4%; Score 1201; DB 1; Length 313;
Best Local Similarity 70.3%; Pred. No. 3.3e-128;
Matches 218; Conservative 41; Mismatches 51; Indels 0; Gaps 0;

QY 2 DTTVPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLKSADLVDFEHGSEA 61
Db 3 DATVPTFDLAELEGLHDEFRCLREKGVFLYLGTLGAEADHASAREIAVDFFDHGTEA 62
QY 62 EKRAVTSVPVPTMRGFTGLSESTAOITNTGSDYSMYSGMTADNLPSPGDFGRITWQ 121
Db 63 EKXAVMTPIIRRGYAGLESESTAOITNTGKYDYSMSYSGMTADNLPSPAEFEKAWED 122
QY 122 YFDRQYTSARAVAREVLRAATGTPDGGVEAFDCEPLLRLFRYPQVPEHRSABEQPLRMA 181
Db 123 YFARMYRASQDVARQVLTSGAEPEVGMDFDCEPLLRLRLRYFPPEVDRVAEEQPLRMA 182
QY 182 PHYDLSMVTLLIQTTPCANGFVSLQAEVGGATDLPYRPDAVLVFCGAIATLVGTGGVAP 241
Db 183 PHYDLSIVTLIHQTPCANGFVSLQVEVDGSDYDIPAPQGAVALVFCGAVATLVADGAIKAP 242
QY 242 RHVVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLAARECGFDVSLDGETATFDWIGG 301
Db 243 KHVVAAPGADKRVGSRTSSVFFLRPNADFTFSPVLAARECGFDVSLDGETATFDWIGG 302
QY 302 YNIRRTSKA 311
Db 303 YINIRKTAA 312

RESULT 3
US-08-592-411-17
; Sequence 17, Application US/08592411
; Patent No. 5726032
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Process for the Efficient Production of
; TITLE OF INVENTION: 7-ADCA via 2-(Carboxyethylthio)acetyl-7-ADCA and
; TITLE OF INVENTION: 3-(Carboxymethylthio)propionyl-7-ADCA
; NUMBER OF SEQUENCES: 17
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,411
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 314 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-592-411-17

Query Match 72.5%; Score 1186.5; DB 1; Length 314;
Best Local Similarity 69.8%; Pred. No. 1.5e-126;
Matches 217; Conservative 41; Mismatches 52; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLKSADLVDFEHGSEA 60
Db 3 DATVPTFDLAELEGLHDEFRCLREKGVFLYLGTLGAEADHASGREIAVDFFDHGTE 62
QY 61 EKRAVTSVPVPTMRGFTGLSESTAOITNTGSDYSMYSGMTADNLPSPGDFGRITW 120
Db 63 EKXAVMTPIIRRGYAGLESESTAOITNTGKYDYSMSYSGMTADNLPSPAEFEKAW 122
QY 121 QYFDRQYTSARAVAREVLRAATGTPDGGVEAFDCEPLLRLFRYPQVPEHRSABEQPLRM 180
Db 123 DYFARMYRASQDVARQVLTSGAEPEVGMDFDCEPLLRLRLRYFPPEVDRVAEEQPLRM 182
QY 181 APHYDLSMVTLLIQTTPCANGFVSLQAEVGGATDLPYRPDAVLVFCGAIATLVGTGGVKA 240
Db 183 APHYDLSIVTLIHQTPCANGFVSLQVEVDGSDYDIPAPQGAVALVFCGAVATLVADGA 242
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLAARECGFDVSLDGETATFDWIGG 300
Db 243 KHVVAAPGADKRVGSRTSSVFFLRPNADFTFSPVLAARECGFDVSLDGETATFDWIGG 302
QY 301 NYNIRRTSKA 311
Db 303 YINIRKTAA 313

RESULT 4
PCT-US95-04801-6
; Sequence 6, Application PC/TUS9504801
; GENERAL INFORMATION:
; APPLICANT: Martin, Juan R.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enguita, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Llaurena, Francisco J.
; APPLICANT: Liras, Paloma
; TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS
; TITLE OF INVENTION: LATE GENES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04801
```

```

US-09-413-231-5
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Wallen III, John W.
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 19179
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3905
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04801-6

Query Match 59.0%; Score 965.5; DB 5; Length 310;
Best Local Similarity 61.2%; Pred. No. 2.1e-101;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAEQGLHODEFRRCCLRDKGLFYLTDGCLTDTLTKSAKDLVIDFPEHGEA 61
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3 DKTVPFVMAELDCSRQDEFREWAR-RGVFLTYCYGATERDHRVATDTAMDFFPAQGTAE 61
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 EKRAVTSVPVPTMRRGFTGLESESTAQITNTGYSYSDSMCYSGMTADNLFFPSGDFGRITWQ 121
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 EQQAVTKVPTMRRGYSALAEASTAQVNTGTYTDSYSGMIGGNLFFPSKEFESVWTD 121
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 YFDRQVTSARAVAREVLRATGTEPDGGEVAFDCEPLLRFRYPFQVPEHRSAAEQPIRMA 181
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 YFDSLRYAAQETARLVLTAAAGTYDGEDLTLLDCDPVLRFLRYFPEVPEHRAAEYEPERMA 181
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 PHYDLSMTVLIQOTPCANGFVLSQAEVGAFTDLPYRDAVLVFCGAIAITVTCGVQKAP 241
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 PHYDLSIITFIHQTPCANGFVLSQAEVDGEMVSLPHVEDAVVLCGAIAITVTCGVQKAP 241
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 RHVVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDMWIGCN 301
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 NHHVSPDASMLKGSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDMWIGCN 301
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 302 YVNI 305
DB |||:
QY 302 YVTM 305
DB |||:

RESULT 5
US-09-413-231-5
Sequence 5, Application US/09413231
Patent No. 6284483
GENERAL INFORMATION:
APPLICANT: Dilley, David R
APPLICANT: Kadyrzhanova, Dina K
APPLICANT: Wang, Zhenyong
APPLICANT: Warner, Toni M
TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and Cephalosporins Under the Control of Bicarbonate
FILE REFERENCE: MSU41-453
CURRENT APPLICATION NUMBER: US/09/413,231
CURRENT FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 329
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: modified IPNS from Streptomyces clavuligerus strain NRRL 3585
NAME/KEY: MUTAGEN
LOCATION: (210)
OTHER INFORMATION: Glu210 in native IPNS modified to Arg

Query Match 10.0%; Score 164; DB 3; Length 329;
Best Local Similarity 23.8%; Pred. No. 4.2e-10;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

QY 5 VPTFSLAEI-----QQGLHODEFRRCCLRDKGLFYLTDGCLTDTLTKSAKDLVIDFPEHG 58
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 59 SEAEK-----RAVTSVPVPTMRRGF-----TGLESESTAQITNTGYSYSDSMCYSGMT---A 106
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 67 TDQEKHDLAIHAYNPDPNPHVRNGYKAVPGKAVESFCYLNPDFGEDHPM-IAAGTPMHE 125
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 107 DNLFPDGS-----FGRITWQYFDRQVTSARAVAREVLRATGTEPDGGEVAFDCE 156
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 126 VNLWPDDEHRHPRFRPCGYYRQMLKSLTVLMRGLALALG-RPEHFFDAALAEQDLSLV 184
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 157 PLLRFRYPFQVPEHRSAAE-QPLRWAPHYDLSMTVLIQOTPCANGFVLSQAEVGAFTDL 215
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 185 SLIRPYLEEYPPVKTGPDGQLLSFRDHLDSMTITVLFPQTQVQN-----LQVETVDGWRDI 240
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 216 PYRPDAVLVFCGAIAITVTCGVQKAPRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSV 275
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 PISENDFVNCGTYMAHVNTDYFPAPNHRV-----KFNVAERLSLFPFLNGCHEAVIE- 293
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 276 PLARECGFDVSLDGETATFQDMI 298
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 294 PFVPE-GASEEVRNEALSYGDYL 315
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
US-09-413-231-9
Sequence 9, Application US/09413231
Patent No. 6284483
GENERAL INFORMATION:
APPLICANT: Dilley, David R
APPLICANT: Kadyrzhanova, Dina K
APPLICANT: Wang, Zhenyong
APPLICANT: Warner, Toni M
TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and Cephalosporins Under the Control of Bicarbonate
FILE REFERENCE: MSU41-453
CURRENT APPLICATION NUMBER: US/09/413,231
CURRENT FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 329
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: modified IPNS from Streptomyces clavuligerus strain NRRL 3585
NAME/KEY: MUTAGEN
LOCATION: (210)
OTHER INFORMATION: Glu210 in native IPNS modified to Arg

Query Match 10.0%; Score 164; DB 3; Length 329;
Best Local Similarity 23.8%; Pred. No. 4.2e-10;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

QY 5 VPTFSLAEI-----QQGLHODEFRRCCLRDKGLFYLTDGCLTDTLTKSAKDLVIDFPEHG 58
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 59 SEAEK-----RAVTSVPVPTMRRGF-----TGLESESTAQITNTGYSYSDSMCYSGMT---A 106
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 67 TDQEKHDLAIHAYNPDPNPHVRNGYKAVPGKAVESFCYLNPDFGEDHPM-IAAGTPMHE 125
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 107 DNLFPDGS-----FGRITWQYFDRQVTSARAVAREVLRATGTEPDGGEVAFDCE 156
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 126 VNLWPDDEHRHPRFRPCGYYRQMLKSLTVLMRGLALALG-RPEHFFDAALAEQDLSLV 184
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 157 PLLRFRYPFQVPEHRSAAE-QPLRWAPHYDLSMTVLIQOTPCANGFVLSQAEVGAFTDL 215
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 185 SLIRPYLEEYPPVKTGPDGQLLSFRDHLDSMTITVLFPQTQVQN-----LQVETVDGWRDI 240
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 216 PYRPDAVLVFCGAIAITVTCGVQKAPRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSV 275
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 PISENDFVNCGTYMAHVNTDYFPAPNHRV-----KFNVAERLSLFPFLNGCHEAVIE- 293
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 276 PLARECGFDVSLDGETATFQDMI 298
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 294 PFVPE-GASEEVRNEALSYGDYL 315
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

US-09-413-231-5
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Wallen III, John W.
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 19179
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3905
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04801-6

Query Match 59.0%; Score 965.5; DB 5; Length 310;
Best Local Similarity 61.2%; Pred. No. 2.1e-101;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAEQGLHODEFRRCCLRDKGLFYLTDGCLTDTLTKSAKDLVIDFPEHGEA 61
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3 DKTVPFVMAELDCSRQDEFREWAR-RGVFLTYCYGATERDHRVATDTAMDFFPAQGTAE 61
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 EKRAVTSVPVPTMRRGFTGLESESTAQITNTGYSYSDSMCYSGMTADNLFFPSGDFGRITWQ 121
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 EQQAVTKVPTMRRGYSALAEASTAQVNTGTYTDSYSGMIGGNLFFPSKEFESVWTD 121
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 YFDRQVTSARAVAREVLRATGTEPDGGEVAFDCEPLLRFRYPFQVPEHRSAAEQPIRMA 181
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 YFDSLRYAAQETARLVLTAAAGTYDGEDLTLLDCDPVLRFLRYFPEVPEHRAAEYEPERMA 181
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 PHYDLSMTVLIQOTPCANGFVLSQAEVGAFTDLPYRDAVLVFCGAIAITVTCGVQKAP 241
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 PHYDLSIITFIHQTPCANGFVLSQAEVDGEMVSLPHVEDAVVLCGAIAITVTCGVQKAP 24
```

Db 126 VNLWPDERRPRPRCEGYRQMLKSLVLMRGLALG-RPEHFFDAALAEQDSSV 184
QY 157 PLIRERYFPQVPEHRSABE-QPLRMAPHYDLSMVTLLIQOTPCANGFVSLQAEVGGAFDTL 215
Db 185 SLIRYPYLEYPPVKTPGDPGQLLSFRDLVDVSMITVLFQTVQN-----LQVETVDGWRDI 240
QY 216 PYRPDAVLVFCGAIAVLTVGGQVKAPRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSV 275
Db 241 PISENDPLVNCGYMAHVINDYPPAHRV-----KFVNAERLSLPFFLNGGHEAVIE- 293
QY 276 PLARECGFVSLDGETATFQDWI 298
Db 294 PFVPE-GASEEVRNEALSVDYL 315

RESULT 7

US-08-379-556A-10
; Sequence 10, Application US/08379556A
; Patent No. 5859329
; GENERAL INFORMATION:
; APPLICANT: HOLTON, TIMOTHY A.
; APPLICANT: KEAM, LISA A.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL
; TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,556A
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9592
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)742-4343
; TELEFAX: (516)742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 335 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-379-556A-10

Query Match 9.9%; Score 162; DB 2; Length 335;
Best Local Similarity 22.8%; Pred. No. 7.3e-10;
Matches 72; Conservative 55; Mismatches 135; Indels 54; Gaps 14;

QY 5 VPTFSLAELOOQGLHQBDFRRCLDKGLFYLTDGGLTDTLKSADLVIDFFEHGSAEKR 64
Db 43 VPIDUSQADNESLVALISKASKDWGIFQVNHGIPSELISKQNVKGKFEFFLPQB-EKE 101
QY 65 AVTSPVPTMRGFTGLSESTAIQITNTGYSYDSCYSGMTADNLF-----PSGDFGRW 119
Db 102 VIAKP-----DGKGVG-----GVTGLQKEVQKKGWDHLEPHIVWPSFINYQFW 148
QY 120 TQ-----YFD--ROYTASR-AVAREVL-----RATGTEPD-----GVFAFLDCEPLLRPR 162
Db 149 PKNPPSYRDTNBEYTSQSLGVANKLLGLLSKGLGLEBEDEVKQALGGEDLIY-----MLKIN 204

QY 163 YFPQVPEHRSABEQLRMAPHYDLSMVTLLIQOTPCANGFVSLQAEVGGAFDTLPYRPDAV 222
Db 205 YTPPCP-----CPALGVAPHTDMSSITL-----VPEVQGLQVFKDGQWYDVAYIPNAL 256
QY 223 LVFCAIAVLTVGGQVKAPRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECG 282
Db 257 ITHIGDQIBLSNGYKSYVHRSTVKNK-----KTRMSWPAFLPEPPEFVG-PIPKLVN 310
QY 283 FVSLDGETATFQDWI 298
Db 311 KODPPKYTKKYKDYV 326

RESULT 8

US-08-379-556A-8
; Sequence 8, Application US/08379556A
; Patent No. 5859329
; GENERAL INFORMATION:
; APPLICANT: HOLTON, TIMOTHY A.
; APPLICANT: KEAM, LISA A.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL
; TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,556A
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9592
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)742-4343
; TELEFAX: (516)742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-379-556A-8

Query Match 9.1%; Score 149; DB 2; Length 333;
Best Local Similarity 21.6%; Pred. No. 2.2e-08;
Matches 71; Conservative 57; Mismatches 126; Indels 74; Gaps 15;

QY 1 MDTT--VPTFSLAELOOQGLHQBDFRRCL-----RDKGLFYLTDGGLTDTLKSADLVIDF 54
Db 39 LITVLEVPAILDS-----LEEDDVVKLVLSASKWGLFOVTNHGIPTEVIEKLVKQVGMF 93
QY 55 FPHGSAEAKRATVSPVPTMRGFTGLSESTAIQITNTGYSYDSCYSGMTADNLPSPGD 114
Db 94 FRAPAE-EKETAKPEGGVEGYGTMLQKEIQGR-----KGWVDHL----- 132
QY 115 FGRINT-----QYF-----DRQYTAARAVAREVL-----RATGTEPDG-----GV 149
Db 133 FHKVWPPSPVINTQWPKPTSYREANETKYLRIVADKLFKCMKGLGLEBEDEVKSCGN 192
QY 150 EAPLDCPEPLLRFRYPQVPEHRSABEQLRMAPHYDLSMVTLLIQOTPCANGFVSLQAEVG 209

Db 193 E---DIVYLLKINYPCCPR-----PDALGVAAHTDLSVITIL-----VPNDVAGLVQVSRD 241
QY 210 GAFTDLPPYRPAVLVFCGAIATLVTTGQVKAPRHHVAAPRRDQIAGSSRTSSVFFLRPNA 269
Db 242 GRWYDVKYIPNALIHVGDOEIMSGEYKAVLHRSVTNKE-----RTRISNPPVFLPEPS 296
QY 270 DFTFSVPLARECGFVSLGSETATFOOM 297
Db 297 DPAVG-PIPKLISDEKPAKYKTKVFSEY 323

RESULT 9
US-09-413-231-10
; Sequence 10, Application US/09413231
; Patent No. 6284483
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/413,231
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; NAME/KEY: MUTAGEN
; LOCATION: (214)
; OTHER INFORMATION: Glu214 in native IPNS modified to Arg
US-09-413-231-10

Query Match 8.7%; Score 143; DB 3; Length 333;
Best Local Similarity 23.2%; Pred. No. 1.1e-07;
Matches 70; Conservative 37; Mismatches 131; Indels 64; Gaps 13;
QY 5 VPTFSLAEQQQLHQD-----EPRCLRDKGLFVLTDCGLTDTTELKSAKDLVIDFFEH 57
Db 10 VPTIDISPL-FGTDPAKAHVARIQNEACRGSGFFVASHGI---DVRRLQDVVNEFHT 65
QY 58 GSAEK-----RAVTSVPVPMRRGFTGLESESTAQITNTGSDYSNCY---SNGTADNL 109
Db 66 MTQEKHDLAIHAYNENNSHVRNGY-----YMARPGKTVESWCYLNPSFGSDHPM 116
QY 110 FPSG-----DFGRITWQYFDROYTASRA-VAREVLRTGT-----EPDG 147
Db 117 IKAGTWMHENVNMPDERHPDFRSFGQYTYREVRFLSKVLLLLGFALALGKPEFFENEV 176
QY 148 GVEAFDLCEPL-LRFYFPQVPE---HRSABEQLRMAPHYDLSMTVLIQQTPCANGFVS 203
Db 177 TEEDTLSCRLMIRYPVLDYPENAIKTPDGTFLSPRDLHVSMTVLFQTEVQN---- 232
QY 204 LQAEVGAFDTLPRPDAVLVFCGAIATLVTTGQVKAPRHHVAAPRRDQIAGSSRTSSVF 263
Db 233 LQVETVDMQSLPTSGENFLNGCTVGLYLTNDYFPAPNHRV-----KYVNAERLSLPF 286
QY 264 FL 265
Db 287 FL 288

RESULT 10
US-08-379-556A-2
; Sequence 2, Application US/08379556A
; Patent No. 5859329
; GENERAL INFORMATION:

; APPLICANT: HOLTON, TIMOTHY A.
; APPLICANT: KEAM, LISA A.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL
; TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,556A
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9592
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)742-4343
; TELEFAX: (516)742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-379-556A-2

Query Match 8.6%; Score 140; DB 2; Length 347;
Best Local Similarity 22.3%; Pred. No. 2.5e-07;
Matches 59; Conservative 43; Mismatches 114; Indels 48; Gaps 9;
QY 27 RKGLFYLTDGLTTELKSAKDLVIDFPEHSEAEKRAVT-SPVPTMERGF-TGLESES 84
Db 77 KEMGIFQLINHIGTDEADLQKVGKEPFEHVHQEELIAKTPTGSDIEGYTSLQKEV 136
QY 85 TQAQITWGSYSDYSCYSMTADNLF-----PSGDFGRW-----TOYFDROYT 128
Db 137 EKG-----KGWVDHLPFKIWPDSAVNRYWPKNPPSYRANEYIGRMRE 181
QY 129 ASRAVAREVLRTATGTEPDGGEVAFDCE--PLLRFYFPQVPEHRSABEQPLRMAPHYDL 186
Db 182 VVDRIFKLSLGLGLEGHMEIEAAGGDEIVLLKINYPCCPR---PDALGVAAHTDM 237
QY 187 SMTVLIQQTPCANGFVSLQAEVGGATDLPYRPDAVLVFCGAIATLVTTGQVKAPRHHVA 246
Db 238 SVITIL-----VPEVNGQLQVFDGHWYDVKYIPNALIHVGDOEILSGKYSVYHRTT 293
QY 247 APRDQIAGSSRTSSVFFLRPNAD 270
Db 294 VNK-----DKTRMSWVPVLEPPSE 312

RESULT 11
US-09-413-231-8
; Sequence 8, Application US/09413231
; Patent No. 6284483
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate

FILE REFERENCE: MSU41-453
CURRENT APPLICATION NUMBER: US/09/413,231
CURRENT FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 321
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
OTHER INFORMATION: from Streptomyces cattleya
NAME/KEY: MUTAGEN
LOCATION: (211)
OTHER INFORMATION: Glu211 in native IPNS modified to Arg
US-09-413-231-8

Query Match 8.4%; Score 137; DB 3; Length 321;
Best Local Similarity 23.1%; Pred. No. 4.8e-07;
Matches 73; Conservative 32; Mismatches 149; Indels 62; Gaps 12;

QY 5 VPTFSLAELOOGLHQDEFR-----RCURDKGLFYLTDGLTDTLTKSAKDLVIDFFEH 57
DB 10 VPTDISPOLFGDPTFRKTSRGRSPARGSGFFYASHHGIDVRLQTSN-----61
QY 58 GSEAKRAVTSPPVTMRRTGFTGLSESTAQ-----ITNTGSDYSYMCY---SMGTADNLF 110
DB 62 -----ESTTMDQSTWRSRYNNNSHVNGYMARPGRETVESWCYLNPSFGEDHPWM 117
QY 111 PSG-----DFGRITWQYDFRQVYASRAV---AREVLRTGTEPDGGV-- 149
DB 118 KAGTPMHEVNVMPDEERHPDFGSGFGQY-HREVSASRRCGASRWRQAGESSNEVTE 176
QY 150 EARLDCEPLLRFRYPQVPE---HRAEEOPLRMAHYLSMVTLLQOTPCANGFVSLQA 206
DB 177 EDTLSAVSMIRYPYLPDPYPEAAIKTGPDRLSFRDHLDSMTILSKTEVQN---LQV 232
QY 207 EVGGAFTDLPYRDAVLVFCGATATLVGGQVKAPRRHHVAAPRRDOIAGSSRTSSVFFELR 266
DB 233 ETVDGHSQSLTSGENFLINGCTYGLTNDYFPAPNHRV-----KVNARLSLPFFLH 286
QY 267 PNADFTFSVPLARECG 282
DB 287 AGQNSVMK-PFTRRTG 301

RESULT 12
US-09-413-231-4
Sequence 4, Application US/09413231
Patent No. 6284483
GENERAL INFORMATION:
APPLICANT: Dilley, David R
APPLICANT: Kadyrzhanova, Dina K
APPLICANT: Wang, Zhenyong
APPLICANT: Warner, Toni M
TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and Cephalosporins Under the Control of Bicarbonate
FILE REFERENCE: MSU41-453
CURRENT APPLICATION NUMBER: US/09/413,231
CURRENT FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 331
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
OTHER INFORMATION: from Penicillium chrysogenum
NAME/KEY: MUTAGEN
LOCATION: (212)
OTHER INFORMATION: Glu212 in native IPNS modified to Arg
US-09-413-231-4

Query Match 8.2%; Score 133.5; DB 3; Length 331;
Best Local Similarity 24.1%; Pred. No. 1.3e-06;
Matches 77; Conservative 39; Mismatches 131; Indels 73; Gaps 14;

QY 27 RDKGLFYLTDG-----LTDTELKSAKDLVIDFF--EHGSE-----60
DB 37 RDTGFFYAVNHGVDVKRLSNKTRPFPFSITDBE---KWDLAIRAYNKEHQDQIRAGYYLS 93
QY 61 -AEKAVTSPVPTMERGFTGLSESESTAQITNTGTSYSDYSCMYCMGTAD--NLFPS-----G 113
DB 94 IPEKKAVES-----FCYL-----NPNFKPDHPLIOSKTPTHEVNVWPDKEKHP 136
QY 114 DFGRITWQYDFRQVYASRAVAREVLRTGTEPDGGVEAF-----LDCEPLLRFRYPQV 168
DB 137 GFREFAEQYVNDVFGLSALLAGYALALGKEEDFFSRHFKKEDALSSVVLIRYPYLNPIP 196
QY 169 EH--RSABE-QPLRMAPHYDLSMVTLLQOTPCANGFVSLQAEGVGAFTDLPYRDPVAVLF 225
DB 197 PAAIKTAEDGTKLSRPHEDVSLITVLYQSDVAN---LQVEMPOGYLDIEADDNAYLVN 252
QY 226 CGAIATLVGGQVKAPRRHHVAAPRRDOIAGSSRTSSVFLRPNADFTFSVPLARECGFDV 285
DB 253 CGSYMAHITNNYYPAPIHRV-----KWNNEERQSLPPFV---NLGFNDIVQWDPSEKDG 304
QY 286 SLDGETATPDQWIGGNVYNI 305
DB 305 KTDQRPISYGDYLNQGLVSL 324

RESULT 13
US-08-553-367A-6
Sequence 6, Application US/08553367A
Patent No. 5939539
GENERAL INFORMATION:
APPLICANT: Theodor LANGE et al.
TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553,367A
FILING DATE: No. 5939539ember 27, 1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 49/FD.4.5M2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO


```
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Arabidopsis thaliana
/ IMMEDIATE SOURCE:
/ CLONE: zu PAT2353
US-08-553-367A-6

Query Match      8.0%; Score 131, DB 2; Length 378;
Best Local Similarity 23.3%; Pred. No. 3e-06;
Matches 60; Conservative 40; Mismatches 130; Indels 28; Gaps 8;

QY 30 GLPYLTDCGLTDELKSAKDLVIDFEHSGEAEKRAVTSVPFT--MRRGFTGLESES--- 84
DB 89 GFPLVNHGVSESLIADAHRLMESFDMPLAGKQAKRPGESCGYASSFTGRFSTKLPW 148
QY 85 ----TAQITNTGYSYDSCYSGMTADNLPSPGDFGRITQYFDROYTASRAVAREVLRA 140
DB 149 KETLSQFQNDNSGSRVTQDYFSDTLGQEF--QFGKVYQDYCEAMSSLSLKIMELLGLS 206
QY 141 TGTEPDGGVEAFDCEPLLRFRYP--QVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCA 198
DB 207 LGVNRDYFRGFPEENDSINRLNHYPCQTP-----DLTGTGPHCDPSSLTILHQDH-V 259
QY 199 NGFVSLQAEVGAFTDLPRPDVILVFCGAIATLVGGQVKAPRHVAAPRRDQIAGSSR 258
DB 260 NG---LQVFVDNQWQSIIRPNKAFVNVIGDTFMALSGNIGFKSCLHRAVVNRE-----SAR 311
QY 259 TSSVFLRPNADFTFSVP 276
DB 312 KSWAFFLCPKDKVKVKKP 329

RESULT 14
US-09-295-306-6
Sequence 6, Application US/09295306
Patent No. 6198021
GENERAL INFORMATION:
APPLICANT: Theodor LANGE et al.
TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/295,306
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/553,367
FILING DATE: No. 6198021ember 27, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 49/DIV-PD4.5M2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
```

```
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Arabidopsis thaliana
/ IMMEDIATE SOURCE:
/ CLONE: zu PAT2353
US-09-295-306-6

Query Match      8.0%; Score 131, DB 3; Length 378;
Best Local Similarity 23.3%; Pred. No. 3e-06;
Matches 60; Conservative 40; Mismatches 130; Indels 28; Gaps 8;

QY 30 GLPYLTDCGLTDELKSAKDLVIDFEHSGEAEKRAVTSVPFT--MRRGFTGLESES--- 84
DB 89 GFPLVNHGVSESLIADAHRLMESFDMPLAGKQAKRPGESCGYASSFTGRFSTKLPW 148
QY 85 ----TAQITNTGYSYDSCYSGMTADNLPSPGDFGRITQYFDROYTASRAVAREVLRA 140
DB 149 KETLSQFQNDNSGSRVTQDYFSDTLGQEF--QFGKVYQDYCEAMSSLSLKIMELLGLS 206
QY 141 TGTEPDGGVEAFDCEPLLRFRYP--QVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCA 198
DB 207 LGVNRDYFRGFPEENDSINRLNHYPCQTP-----DLTGTGPHCDPSSLTILHQDH-V 259
QY 199 NGFVSLQAEVGAFTDLPRPDVILVFCGAIATLVGGQVKAPRHVAAPRRDQIAGSSR 258
DB 260 NG---LQVFVDNQWQSIIRPNKAFVNVIGDTFMALSGNIGFKSCLHRAVVNRE-----SAR 311
QY 259 TSSVFLRPNADFTFSVP 276
DB 312 KSWAFFLCPKDKVKVKKP 329

RESULT 15
US-09-734-719-6
Sequence 6, Application US/09734719
Patent No. 6455675
GENERAL INFORMATION:
APPLICANT: Theodor LANGE et al.
TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/734,719
FILING DATE: 13-Dec-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/295,306
FILING DATE: April 21, 1999
APPLICATION NUMBER: 08/553,367
FILING DATE: No. 6455675ember 27, 1995
APPLICATION NUMBER: PCT/EP94/01664
FILING DATE: May 24, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 2000-1678/LC/01784
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX: <Unknown>
```

```
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; IMMEDIATE SOURCE:
; CLONE: zu PAT2353
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-734-719-6

Query Match      8.0%; Score 131; DB 4; Length 378;
Best Local Similarity 23.3%; Pred. No. 3e-06;
Matches 60; Conservative 40; Mismatches 130; Indels 28; Gaps 8;

QY 30 GLPYLTDCGLTDELKSAKDLVIDFPEHGSEAEKRAVTSVPPT--MRGFTGLESES--- 84
Db 89 GFLVNVHGVSESLIADAHRLMESFFDMPLAGKQAKRPGESCGYASSFTGRFSTKLPW 148
QY 85 ---TAQITNTGYSYSCYSMGTADNLPPSGDFGR1WTQYFDROYTASRAVAREVLRA 140
Db 149 KETLSFQFSNDNSGSRVQDYFSDTLQGEFE--QFGKVYQDYCEAMSSLSLKIMELLGLS 206
QY 141 TGTEPDGGVEAFIDCEPLARFRYP--QVPEHRSABEEQPLRMAPHYDLSMVTLIQOTPCA 198
Db 207 LGVNRDYFRGFPEENDSIMELNHVPCQTP-----DLTLGTGPHCDPSSLTILHQDH-V 259
QY 199 NGFVSLQAEVGGFTDLPYRPDAVLVFCGAIATLVGGQVKAPRHHVAAPRRDQIAGSSR 258
Db 260 NG---LQVFVDNQWQ3IRPNPKAFVNVNIGDTFMALNSGIFKSLHRAVVNRE-----SAR 311
QY 259 TSSVFFLRPNADFTFSVP 276
Db 312 KSMAFFLCPKKDKVVKPP 329
```

Search completed: February 3, 2004, 17:43:00
Job time : 22 secs

Qy	MDTVPPTFSIAELQQGLHQDEFRRCIRDKGLFYLTDCGLTDTELKSADLVIDFFHGSE	60
Dβ	MDTVPPTFSIAELQQGLHQDEFRRCIRDKGLFYLTDCGLTDTELKSADLVIDFFHGSE	60

QY 61 AEKRAVTSVPVPTMRRGFTGLESESTAQITNTGSDYSCMYCMGTADNLPFGDGFRIWT 120
DB 61 AEKRAVTSVPVPTMRRGFTGLESESTAQITNTGSDYSCMYCMGTADNLPFGDGFRIWT 120
QY 121 QYFDROYTASRAVAREVLATGTEPPGGVEAFDCEPLLRFYFPQVPEHRSABEQPLRM 180
DB 121 QYFDROYTASRAVAREVLATGTEPPGGVEAFDCEPLLRFYFPQVPEHRSABEQPLRM 180
QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDAVLVFCGAIATLVTTGGQVKA 240
DB 181 RPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDAVLVFCGAIATLVTTGGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311
RESULT 2
US-10-369-493-12239
; Sequence 12239, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12239
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-12239

Query Match 10.3%; Score 168; DB 12; Length 366;
Best Local Similarity 24.7%; Pred. No. 9.6e-10;
Matches 85; Conservative 41; Mismatches 158; Indels 60; Gaps 14;
QY 1 MDVTTFSLAELOOGLHQ-----DBFRCLRDKGLFYLTDCGLTDTLTKSAKDLVIDFF 55
DB 28 MPRIVPVLDSRLUEQASERTFLADLRASRDIGFFYLAGHIGSWAEISEVLTASRQFF 87
QY 56 EHGSEAEKRAVTSVPVPTMRRGFTGLESESTAQITNTGSDYSCMYCMGTADNLPFGSGDF 115
DB 88 AL-PEADKLAIEVKSQFQGYTRAGGELTK-----GREDWREQLDIGVERQAIAGPG 140
QY 116 GRITWQ-YFDRQYTA-----RAVAREVLATGTEPPGGVEAFDCEPL 159
DB 141 TPANTRLOQGNQWPAALPDLKALLAWQSKVTAVARLLKFAQSLDQPEDAF---DPI- 196
QY 160 RPYFPQVPEHR-SABEQPLR-----MAPHYDLSMVTLIQOTPCANGFVSLQAEVGG 210
DB 197 ----YSEFNRHKMIVRYFGDRTGGDQGVGAHKGGLTLILQ-----DDNKGLOVDYDG 248
QY 211 AFTDLPYRDAVLVFCGAIATLVTTGGQVKAAPRHVAAPRRDQIAGSSRTSSVFFLRPNAD 270
DB 249 SWVDVDP1PGTLLVNVNIGELLELASNGYLRTAVHRVTTTP-----AGVERISVPVFFSARLD 304
QY 271 FT-----FSVPLARECGFDVSLDGETATFQDWIGGNVNIIRRTS 309
DB 305 ATPLGLSLSELAQAARGPAS-DFDNPFLRD-VGTNVLKSLRLS 346

RESULT 3
US-09-924-841-5
; Sequence 5, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/924,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Streptomyces clavuligerus
; NAME/KEY: MUTAGEN
; LOCATION: (210)
; OTHER INFORMATION: Glu210 in native IPNS modified to Arg
US-09-924-841-5

Query Match 10.0%; Score 164; DB 10; Length 329;
Best Local Similarity 23.8%; Pred. No. 2.3e-09;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;
QY 5 VPTFSLAEI-----QQGLHQDEFRRCLRDKGLFYLTDCGLTDTLTKSAKDLVIDFFHG 58
DB 10 VPTIDISPLFGTDAAAKKRAVEIHGACRGSGFFVATNHGV---DVQQLQDVVNEFHGAM 66
QY 59 SEAEK-----RAVTSVPVPTMRRGFT-----TGLSESTAQITNTGSDYSCMYCMGT---A 106
DB 67 TDOEKHDLAIHAYNPDPNPHVRNGYKAVPRKAVESFCYLNPDFGDEHFM-IAAGTPMHE 125
QY 107 DNLFPDGD-----FGRIWQYFDRQYTSRAVAREVLATGTEPPGGVEA-----FLDCE 156
DB 126 VNLWPDDEERHPRFPCGEGYQMLKLSIVLMRGLALALG-RPEHFFDAALAEQDSLSV 184
QY 157 PLLRFYFPQVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFD 215
DB 185 SLIRPYLSEYPPVKTGPDGQLLSPRDHLDSMITVLFQTVQN-----LQVETVDCWRDI 240
QY 216 PYRPAVLVFCGAIATLVTTGGQVKAAPRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFV 275
DB 241 PTSENDPLVNCGTYMAHVNTDYPAPNHRV-----KFYNAERLSLPFFLNGHGAIE- 293
QY 276 PLARECGFDVSLDGETATFQDWI 298
DB 294 PFVPE-GASEEVRNEALSIGDYL 315

RESULT 4
US-09-924-841-9
; Sequence 9, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/924,841

1 ; CURRENT FILING DATE: 2001-08-08
2 ; PRIOR APPLICATION NUMBER: US/09/413,231
3 ; PRIOR FILING DATE: 1999-10-06
4 ; NUMBER OF SEQ ID NOS: 18
5 ; SOFTWARE: PatentIn Ver. 2.0
6 ; SEQ ID NO 9
7 ; LENGTH: 329
8 ; TYPE: PRT
9 ; ORGANISM: Artificial Sequence
10 ; FEATURE:
11 ; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
12 ; OTHER INFORMATION: from Streptomyces clavuligerus strain NRRL 3585
13 ; NAME/KEY: MUTAGEN
14 ; LOCATION: (210)
15 ; OTHER INFORMATION: Glu210 in native IPNS modified to Arg
16 ; US-09-924-841-9
17
18 Query Match 10.0%; Score 164; DB 10; Length 329;
19 Best Local Similarity 23.8%; Pred. No. 2.3e-09;
20 Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;
21
22 QY 5 VPTFSLAEL-----QQGLHQDFRCLRDKGLFYLTDCGLTDTTELKSAKDLVIDFHEG 58
23 Db 10 VPTIDISPLGTTAAAKRVAEIHGACRSGGFYATNHCV---DVQQLQDVVNEFHGM 66
24
25 QY 59 SEAEK-----RAVTSVPVTTMRGF---TGLESESTAQITNTGSYSDYSCYSGMT---A 106
26 Db 67 TDQEKHDLAIHAYNPDPNPHVNGYKAVPCRKAVESFCYLNLPDGEDHPM-IAAGTPMHE 125
27
28 QY 107 DNUFPSPGD-----FGRITWQYFDROYTASRAVAREVLRAVTEPDGGVEA-----FLDCE 156
29 Db 126 VNLWPDDEERHPRFPCEGYRQMLKSLVLMRGLALALG-RPEHFFDALAEQDSLSSV 184
30
31 QY 157 PLIRFYFPQVPEHRAAE-QPLRMAPHYDLSMTVLIQQTCCANGFVSLQAEVGGAFD 215
32 Db 185 SLIRFYPLEEYPPVKTPGDCQLLSFRDHLQVSMITVLFQVQN---LQVETVQGRDI 240
33
34 QY 216 PYRPDAVLVFCGAIATLVTTGGQVKAPRHVVAAPRRDQIAGSSRTSSVFFLRPNADFTFSV 275
35 Db 241 PTSENDFVNCGYMAHVNTDYPAPNHRV-----KFVNAERLSLPFLNGGHEAVIE- 293
36
37 QY 276 PLARECGFDVSLDGETATFODWI 298
38 Db 294 PFVPE-GASEBEVRNEALS YGDYL 315
39
40 RESULT 5
41 US-10-369-493-14043
42 *; Sequence 14043, Application US/10369493
43 ; Publication No. US20030233675A1
44 ; GENERAL INFORMATION:
45 ; APPLICANT: Cao, Yongwei
46 ; APPLICANT: Hinkle, Gregory J.
47 ; APPLICANT: Slater, Steven C.
48 ; APPLICANT: Goldman, Barry S.
49 ; APPLICANT: Chen, Xianfeng
50 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
51 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
52 ; FILE REFERENCE: 38-10(52052)B
53 ; CURRENT APPLICATION NUMBER: US/10/369,493
54 ; CURRENT FILING DATE: 2003-02-28
55 ; PRIOR APPLICATION NUMBER: US 60/360,039
56 ; PRIOR FILING DATE: 2002-02-21
57 ; NUMBER OF SEQ ID NOS: 47374
58 ; SEQ ID NO 14043
59 ; LENGTH: 327
60 ; TYPE: PRT
61 ; ORGANISM: Pseudomonas fluorescens
62 ; US-10-369-493-14043
63
64 Query Match 9.1%; Score 148.5; DB 12; Length 327;
65 Best Local Similarity 23.8%; Pred. No. 1.2e-07;
66 Matches 80; Conservative 43; Mismatches 162; Indels 51; Gaps 13;

QY 5 VPTFSLAEL-----QQGLHQDFRCLRDKGLFYLTDCGLTDTTELKSAKDLVIDFHEGS 59
Db 1 LPILDLSLLDGTSPSQROAFDLDLRAARDVGFYLTGHGIDAGLLREQVDYARQFAL-P 59
QY 60 EAEKRAVTSVPVTTMRGFTGLESESTAQITNTGSYSDYSCYSGMTADNLFPSPGDFGRW 119
Db 60 DSEKNAVGMINSFHRGYNRAASEITR-----GQPDQREQFDLGAERDVLPLNADSP 113
QY 120 TO-----YFDRQVTASRAVAREVLRAVG---TEPDGGVEAFDCEPLL 159
Db 114 ARLOGPNQWPGALPQLKPLLLDQWQWNTOMSLR-LLRAFAQALSLEDAFDRLYLGKPN 172
QY 160 RFRYPQVPEHRSABEQPLRMAPHYDLSMTVLIQQTCCANGFVSLQAEV-GGAFTDLPYR 218
Db 173 HIKLMRYPGQASTASNQ--GUGAHKDSGFLSFLQDQQA---GLQVEIEEGRWIDALPR 226
QY 219 PRAVLVFCGAIATLVTTGGQVKAPRHVVAAPRRDQIAGSSRTSSVFFLRPNAD-----FTF 273
Db 227 DNTLVVNIIGELLELATNGYLRAVTVHRVVSPP-----VGSERLSIAFFLGAQLDVA 282
QY 274 SVPLARECGFDVSLDGETATFODWIGNVNINRRTS 309
Db 283 PTALLREARCPAS-DPLNPLFRD-VGVNLYKGLR 316
RESULT 6
US-09-924-841-10
; Sequence 10, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Tom M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/924,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from streptomyces anulatus (S. lipmanii)
; NAME/KEY: MUTAGEN
; LOCATION: (214)
; OTHER INFORMATION: Glu214 in native IPNS modified to Arg
; US-09-924-841-10
Query Match 8.7%; Score 143; DB 10; Length 333;
Best Local Similarity 23.2%; Pred. No. 5.1e-07;
Matches 70; Conservative 37; Mismatches 131; Indels 64; Gaps 13;
QY 5 VPTFSLAELQQGLHQD-----EPRCLRDKGLFYLTDCGLTDTTELKSAKDLVIDFHEH 57
Db 10 VPTIDISPL-FGTDPAKAKHVARQINEACRSGGFYASHHGI---DVRRLQDVVNEPHRT 65
QY 58 GSEAEK-----RAVTSVPVTTMRGFTGLESESTAQITNTGSYSDYSCY---SMGTADNL 109
Db 66 MTDQEKHDLAIHAYNNNSHVRNGY-----YMARPGKTKVESWCYLNLPSPFGEDHPM 116
QY 110 FPSG-----DFGRITWQYFDROYTASRA-VAREVLRAVTEG-----EPDG 147
Db 117 IKAGTPMHEVNWVWPDDEERHPRFPFSGEQYVREVRSLKVLRLGFPALALGKPEFFENEV 176
QY 148 GVEAFDCEPL-LRFRYFPQVPE---HRSAEQPLRMAPHYDLSMTVLIQQTCCANGFVS 203

Db 177 TEBDTLSCRLMRYPLDYPPEAAIKTGDPGRFLSFRDHLDSVMTLVLFQTEVQN----232
QY 204 LQEVGGAFTDLPRDAVLVFCGATATLVGGQVKAAPRHVVAAPRRDQIAGSSRTSSVP 263
Db 233 LQVETVDGWSLPTSGENFLNGTYLGLVLTNDYFPAPNHRV-----KYVNAERLSLFP 286
QY 264 FL 265
Db 287 FL 288

RESULT 7
US-10-136-444-8
; Sequence 8, Application US/10136444
; Publication No. US20030101477A1
; GENERAL INFORMATION:
; APPLICANT: COLLIVER, Steven P.
; APPLICANT: HUGHES, Stephen G.
; APPLICANT: MUIR, Shelagh R.
; APPLICANT: TUNEN van, Adrianus J.
; APPLICANT: VERHOEYEN, Martine E.
; TITLE OF INVENTION: A PROCESS FOR INCREASING THE FLAVONOID CONTENT OF A PLANT AND PLA
; FILE REFERENCE: 056159-5106
; CURRENT APPLICATION NUMBER: US/10/136,444
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: EP 01304009.2
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Petunia hybrida
US-10-136-444-8

Query Match 8.6%; Score 140; DB 15; Length 359;
Best Local Similarity 22.3%; Pred. No. 1.2e-06;
Matches 59; Conservative 43; Mismatches 114; Indels 48; Gaps 9;
QY 27 RDKGLFYLDGTLDTLSAKDLVIDDFEHGSAEKRAVT-SPVPTMRGPF-TGLESES 84
Db 88 KEWGIFQLNHGIPDEAIDLQKVGKEFFEHVQEEKELIAKTPGSDNIEGYGTSIQKEV 147
QY 85 TAQITWGSYDSYCMYSGTADNLF-----PSGDFGRW-----TQYFDRQYT 128
Db 148 EKG-----KGVVDHLFKIWPSSAVNRYWPKNPPSYREANEYGRMRE 192
QY 129 ASRAVAREVLRAATGTPDGGVEAFDCE--PLLRFRYFQVPEHRSABEQPLRMAPHYDL 186
Db 193 VVDRIFKISLGLGLEGHEMIEAAGDEIVYLLKINYPCCPR-----PDLALGVVAHTDM 248
QY 187 SMVTLLIQTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGATATLVGGQVKAAPRHVVA 246
Db 249 SVITIL----VPNEVQLQVFKDGHVYDVKYIINALIVHIGDQVEILSNGKYKSVYHRTT 304
QY 247 APRDQIAGSSRTSSVFFLRPNAD 270
Db 305 VNK-----DKTRMSWPVLEPPSE 323

RESULT 8
US-10-431-273-88
; Sequence 88, Application US/10431273
; Publication No. US20030237108A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Glenn, Matthew
; APPLICANT: No. US20030237108A1riss, Michael Geoffrey
; APPLICANT: Saulsbury, Keith Martin
; APPLICANT: Hall, Claire

; APPLICANT: Forster, Richard L. S.
; TITLE OF INVENTION: Compositions isolated from forage
; FILE REFERENCE: 11000.1069U
; CURRENT APPLICATION NUMBER: US/10/431,273
; CURRENT FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Lolium perenne
US-10-431-273-88

Query Match 8.5%; Score 139.5; DB 12; Length 380;
Best Local Similarity 23.3%; Pred. No. 1.5e-06;
Matches 70; Conservative 43; Mismatches 140; Indels 47; Gaps 11;
QY 1 MDTVTFTSLAELOQQLHQDEPR-----RCLRDKGLFYLDGTLDTLSAKDLV 51
Db 60 LOTSIPVIDLGEVLAAT-ADEGRMRQIMEAVAAACREWGFFQVNVHGVAPELMHAREAW 118
QY 52 IDFFHGSAEKRAVTSVPVPTMRGFTGLESESTQITNTGYSYCMYSGTADNLP 111
Db 119 RGFRLPITAKQYANLP-----RTYEGYGSRVGVKGGPLDWGDY---YFLHLAPDAGK 170
QY 112 SGDFGRWITQYFDRQYTASRAVAREVL-----ATCTPDGGVEAF--LDCEPL 158
Db 171 SPD--KYWPTNPAICKDVSEYGREVIRCELLMKMSASLGLEATRFQEAFGGSGCVC 228
QY 159 LRFRYFQVPEHRSABEQPLRMAPHYDLSMVTLLIQTPCANGFVSLQAEVGAFTDLPYR 218
Db 229 LRANYPCPQ----PDLTLGLSAHSDPGVLTIVLADEHVRGLQVRRAD-GEWVTYQPAR 283
QY 219 PDAVLVFCGATATLVGGQVKAAPRHVVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLA 278
Db 284 HDAFTVNVGDQIQLLSNMYKSVHRVMVNAKEE-----RISLALFYNPRGD----VPIA 334

RESULT 9
US-09-924-841-8
; Sequence 8, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/924,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Streptomyces cattleya
; NAME/KEY: MUTAGEN
; LOCATION: (211)
; OTHER INFORMATION: Glu211 in native IPNS modified to Arg
US-09-924-841-8

Query Match 8.4%; Score 137; DB 10; Length 321;
Best Local Similarity 23.1%; Pred. No. 2.3e-06;
Matches 73; Conservative 32; Mismatches 149; Indels 62; Gaps 12;

Query Match	8.3%;	Score 136.5;	DB 12;	Length 259;
Best Local Similarity	25.2%;	Pred. No. 1.8e-06;		
Matches	67;	Conservative 41;	Mismatches 111;	Indels 47; Gaps 11;
Qy	30	GLFVLTGGLTDTLEKSAKDLVIPPFGHGSABEKAQVTSVPVTRRGFTGLESESTAQIT	89	
Db	11	GFAVVRDHGVQPELIERAWRLTKAFFDL-PEEEKRAHFIPGGGARGVTPPKTE-----IA	65	
Qy	90	NTGSYSYSMCYSWG-----TADNLFPSGDFGRITWQYFDRQYATASAVAREV	137	
Db	66	KGATHDLKKEFWHIGROLAAGHRFADVMAPNIWTPRPE--FRETFTLFAAFDAAGDKL	123	
Qy	138	LRAT-----GTEPDGGEAFLLDCEPLLRFRYPFPQPEHRSAEQRLRMAPHVDLSMVTLIQ	193	
Db	124	LSAVARYLGLAPDWFDTAVKDGNVLRLLHYPPV-----ADAPEVRAGAHEDINLIITLL-	178	
Qy	194	QTPCANGFVSLQAEVGG-AFTD-----LPYRP--DAVLVFCGAJATLVTGGQVKAPRHV	245	
Db	179	-----LGAEEAGLELLDRGKMLAVKPPPEGAMVINVDMLQRLTNHVLFSPTTHRV	228	
Qy	246	AAPRRDQIAGSSRTSSVFFFLRPNADF	271	
Db	229	VNPPAER-RGHSRYSMPPFFLHPADPF	253	

```

RESULT 12
US-10-369-493-7649
; Sequence 7649, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MITC
; TITLE OF INVENTION: PLANTS WITH IMPR
; FILE REFERENCE: 38-10(52052)B

```


Search completed: February 3, 2004, 17:42:26
Job time : 34 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:35:49 ; Search time 175 Seconds
(without alignments)
1617.056 Million cell updates/sec

Title: US-09-582-486-1

Perfect score: 1637

Sequence: 1 MDTVTFTSLAELOQLHOD.....ATFDWIGNYNVRTSKA 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*

2: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*

3: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*

4: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

5: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

6: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

7: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

8: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

9: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

10: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

11: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

12: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

13: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

14: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

15: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

16: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

17: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

18: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

19: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

20: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

21: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

22: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

23: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

24: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

25: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

26: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

27: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

28: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

29: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

30: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

31: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

32: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1637	100.0	311	19	US-09-582-486-1

Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-582-486-1

Sequence 1, Application US/09582486

GENERAL INFORMATION:

APPLICANT: SCHOFIELD, Christopher J.

APPLICANT: BALDWIN, Jack E.

APPLICANT: LLOYD, Matthew D.

APPLICANT: HARLOS, Karl

APPLICANT: ANDERSSON, Inger

APPLICANT: TERWISSCHA VAN SCHELTINGA, Anke S.

APPLICANT: VALEGARD, Karin

APPLICANT: RAMASWAMY, S.

TITLE OF INVENTION: MODIFIED DEACETOXYCEPHALOSPORIN C SYNTHASE (DAOS) AND X-RAY STRUCTURE

FILE REFERENCE: 08004624

CURRENT APPLICATION NUMBER: US/09/582,486

CURRENT FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: 9727370.0

PRIOR FILING DATE: 1997-12-24

PRIOR APPLICATION NUMBER: 9813644.3

PRIOR FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 1

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 1

Sequence 32088, A

Sequence 11094, A

Sequence 18, Appli

Sequence 121286, A

Sequence 6, Appli

Sequence 6, Appli

Sequence 10913, A

Sequence 10914, A

Sequence 4795, Ap

Sequence 94436, A

Sequence 4795, Ap

Sequence 89857, A

Sequence 9085, Ap

Sequence 12239, A

Sequence 12239, A

Sequence 89855, A

Sequence 5, Appli

Sequence 9, Appli

Sequence 9082, Ap

Sequence 89847, A

Sequence 52994, A

Sequence 58493, A

Sequence 146275, A

Sequence 51128, A

Sequence 63175, A

Sequence 77410, A

Sequence 41198, A

Sequence 59894, A

Sequence 17598, A

Sequence 136598, A

Sequence 52863, A

Sequence 62424, A

Sequence 53454, A

Sequence 72059, A

Sequence 106840, A

Sequence 50491, A

Sequence 50490, A

Sequence 15162, A

Sequence 18920, A

Sequence 18920, A

Sequence 14043, A

Sequence 14043, A

Sequence 148589, A

311 22 US-09-791-537-32088

311 22 US-09-791-537-11094

311 24 US-09-924-841-18

314 22 US-09-791-537-121286

310 6 US-08-233-605-6

310 11 US-08-727-468-6

310 22 US-09-791-537-10913

318 22 US-09-791-537-10914

319 22 US-09-791-537-4795

332 22 US-09-791-537-94436

332 22 US-09-791-537-4795

332 22 US-09-791-537-89857

332 22 US-09-791-537-9085

366 29 US-10-369-493-12239

366 32 US-60-360-039-12239

329 22 US-09-791-537-89855

329 24 US-09-924-841-5

329 24 US-09-924-841-9

326 22 US-09-791-537-9082

326 22 US-09-791-537-89847

342 28 US-10-219-999-52994

375 30 US-10-425-114-58493

375 30 US-10-425-114-58493

366 30 US-10-437-963-146275

352 28 US-10-219-999-51128

352 30 US-10-425-114-63175

339 21 US-09-708-427-77410

342 21 US-09-708-427-77409

342 28 US-10-219-999-41198

342 28 US-10-219-999-59894

342 32 US-60-324-109-17598

350 30 US-10-437-963-136698

371 28 US-10-219-999-52863

371 30 US-10-425-114-62424

372 30 US-10-425-114-53454

372 30 US-10-425-114-72059

331 22 US-09-791-537-106840

338 21 US-09-708-427-50491

341 21 US-09-708-427-50490

327 24 US-09-902-540-15162

440 21 US-09-733-089-18920

440 21 US-09-816-660-18920

327 29 US-10-369-493-14043

327 32 US-60-360-039-14043

389 30 US-10-437-963-148589

99.5

99.4

99.2

72.5

59.0

59.0

59.0

57.4

57.2

56.6

56.6

10.3

10.3

10.0

10.0

10.0

9.9

9.9

9.9

9.5

9.5

9.4

9.4

9.4

9.3

9.3

9.3

9.3

9.3

9.3

9.2

9.2

9.1

9.1

9.1

9.1

9.1

9.1

9.1

```
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-582-486-1

Query Match      100.0%; Score 1637; DB 19; Length 311;
Best Local Similarity 100.0%; Pred. No. 8.5e-179;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTVTFTSLAEIQOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVDFFEHSGE 60
DB 1 MDTVTFTSLAEIQOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVDFFEHSGE 60
QY 61 AEKRAVTSVPVTRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVTRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
QY 121 QYFDROYTASRAVAREVLRATGTEPDGGVEAFDCEPFLFRFYFPQVPEHRSAAEQPLRM 180
DB 121 QYFDROYTASRAVAREVLRATGTEPDGGVEAFDCEPFLFRFYFPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240
QY 241 PRHHVAAPRRDOIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFDQWIGG 300
DB 241 PRHHVAAPRRDOIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFDQWIGG 300
QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311

Query Match      99.5%; Score 1629; DB 22; Length 311;
Best Local Similarity 99.7%; Pred. No. 7.1e-178;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTVTFTSLAEIQOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVDFFEHSGE 60
DB 1 MDTVTFTSLAEIQOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVDFFEHSGE 60
QY 61 AEKRAVTSVPVTRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVTRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
QY 121 QYFDROYTASRAVAREVLRATGTEPDGGVEAFDCEPFLFRFYFPQVPEHRSAAEQPLRM 180
DB 121 QYFDROYTASRAVAREVLRATGTEPDGGVEAFDCEPFLFRFYFPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240

; LENGTH: 311
; TYPE: PRT
; ORGANISM: pdb 1DCS
US-09-791-537-32088

Query Match      99.4%; Score 1627; DB 22; Length 311;
Best Local Similarity 99.4%; Pred. No. 1.2e-177;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTVTFTSLAEIQOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVDFFEHSGE 60
DB 1 MDTVTFTSLAEIQOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVDFFEHSGE 60
QY 61 AEKRAVTSVPVTRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVTRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
QY 121 QYFDROYTASRAVAREVLRATGTEPDGGVEAFDCEPFLFRFYFPQVPEHRSAAEQPLRM 180
DB 121 QYFDROYTASRAVAREVLRATGTEPDGGVEAFDCEPFLFRFYFPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240

; LENGTH: 311
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-791-537-11094

Query Match      99.4%; Score 1627; DB 22; Length 311;
Best Local Similarity 99.4%; Pred. No. 1.2e-177;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTVTFTSLAEIQOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVDFFEHSGE 60
DB 1 MDTVTFTSLAEIQOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVDFFEHSGE 60
QY 61 AEKRAVTSVPVTRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVTRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
QY 121 QYFDROYTASRAVAREVLRATGTEPDGGVEAFDCEPFLFRFYFPQVPEHRSAAEQPLRM 180
DB 121 QYFDROYTASRAVAREVLRATGTEPDGGVEAFDCEPFLFRFYFPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240

; LENGTH: 311
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-791-537-11094
```

```
QY 241 PRHHVAAPRRDOIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFDQWIGG 300
DB 241 PRHHVAAPRRDOIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFDQWIGG 300
QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311

RESULT 3
US-09-791-537-11094
; Sequence 11094, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11094
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-791-537-11094

Query Match      99.4%; Score 1627; DB 22; Length 311;
Best Local Similarity 99.4%; Pred. No. 1.2e-177;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTVTFTSLAEIQOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVDFFEHSGE 60
DB 1 MDTVTFTSLAEIQOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVDFFEHSGE 60
QY 61 AEKRAVTSVPVTRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVTRRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRITW 120
QY 121 QYFDROYTASRAVAREVLRATGTEPDGGVEAFDCEPFLFRFYFPQVPEHRSAAEQPLRM 180
DB 121 QYFDROYTASRAVAREVLRATGTEPDGGVEAFDCEPFLFRFYFPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240

; LENGTH: 311
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-791-537-11094
```

```
US-09-924-841-18
; Sequence 18, Application US/09924841
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/924,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
```

PRIOR FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 18
LENGTH: 311
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: modified DAOCs
OTHER INFORMATION: from Streptomyces clavuligerus
NAME/KEY: MUTAGEN
LOCATION: (181)
OTHER INFORMATION: Ala181 in native DAOCs modified to Arg
US-09-924-841-18

Query Match 99.2%; Score 1624; DB 24; Length 311;
Best Local Similarity 99.4%; Pred. No. 2.7e-177;
Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDTTPTFSLAEIQOGLHQDEFRRCLRDKGLFYLTDCGLTDTLTKSAKOLVIDFFHGSE 60
DB 1 MDTTPTFSLAEIQOGLHQDEFRRCLRDKGLFYLTDCGLTDTLTKSAKOLVIDFFHGSE 60

QY 61 AEKRAVTSPTVTRRGFTGLESTTAQITNTGYSYDSCYSGMTADNLPSPGDFGRWT 120
DB 61 AEKRAVTSPTVTRRGFTGLESTTAQITNTGYSYDSCYSGMTADNLPSPGDFGRWT 120

QY 121 QYEDROYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
DB 121 QYEDROYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180

QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRDPDAVLVFCGAIATLTGQVKA 240
DB 181 RPHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRDPDAVLVFCGAIATLTGQVKA 240

QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300

QY 301 NYNIRRTSKA 311
DB 301 NYNIRRTSKA 311

RESULT 5
US-09-791-537-121286
; Sequence 121286, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 121286
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Streptomyces lactamdurans
US-09-791-537-121286

Query Match 72.5%; Score 1186.5; DB 22; Length 314;
Best Local Similarity 69.8%; Pred. No. 7.1e-127;
Matches 217; Conservative 41; Mismatches 52; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAEIQOGLHQDEFRRCLRDKGLFYLTDCGLTDTLTKSAKOLVIDFFHGSE 60
DB 3 DATVPTDLAEIQLGREGHREHCLREKGVFLKGTGLPAEADHASGREIAVDFFHGTE 62

QY 61 AEKRAVTSPTVTRRGFTGLESTTAQITNTGYSYDSCYSGMTADNLPSPGDFGRWT 120

DB 63 AEKKAVMPTPIPTIRRGYAGLESESTAQITNTGKTDYSMSYSGMTADNLPSPSAEFSKAW 122

QY 121 QYEDROYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180

DB 123 DIFARMYRASQDVARQVLTSGAEPEVGMDAFDCEPLLRFRYPQVPEHRSAAEQPLRM 182

QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRDPDAVLVFCGAIATLTGQVKA 240

DB 183 APHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRDPDAVLVFCGAIATLTGQVKA 242

QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300

DB 243 PKHVAAPGADKRVGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 302

QY 301 NYNIRRTSKA 311

DB 303 NYNIRRTSKA 313

RESULT 6
US-08-233-605-6
; Sequence 6, Application US/08233605
; GENERAL INFORMATION:
; APPLICANT: Martin, Juan P.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enquita, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Llaras, Paloma
; APPLICANT: Llaras, Paloma
; TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/233,605
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen III, John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19179
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-233-605-6

Query Match 59.0%; Score 965.5; DB 6; Length 310;
Best Local Similarity 61.2%; Pred. No. 2e-101;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAEIQOGLHQDEFRRCLRDKGLFYLTDCGLTDTLTKSAKOLVIDFFHGSEA 61
DB 3 DKTVPFVMAELRDSGRQDEFREWAR-RGVFYLTYGATERDHRVATDTAMDFFAQGTAE 61

QY 62 EKRAVTSVPVPTMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLFPSPGDFGRITWQ 121
Db 62 EKQAVTTKVPVPTMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLFPSPGDFGRITWQ 121
QY 122 YFDRQYTSRAVARAVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRMA 181
Db 122 YFDSLRYRAAQETARLVLTAAAGTYDGEDLTLLDCDPVLRVFPPEHRAAAEYEPERRMA 181
QY 182 PHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVTTGGQVAP 241
Db 182 PHYDLSIIITFIHQTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVTTGGQVAP 241
QY 242 RHVVAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGN 301
Db 242 NHVWSPDASMLKGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGN 301
QY 302 YVNI 305
Db 302 YVTM 305

RESULT 7

US-08-727-468-6
; Sequence 6, Application US/08727468
; GENERAL INFORMATION:
; APPLICANT: Martin, Juan P.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enguita, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Llarena, Francisco J.
; APPLICANT: Liras, Paloma
; TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS
; TITLE OF INVENTION: LATE GENES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/233,605
; FILING DATE: 22-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen III, John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-727-468-6

Query Match 59.0%; Score 965.5; DB 11; Length 310;
Best Local Similarity 61.2%; Pred. No. 2e-101;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAELOQGLHODEFRRCRLDKGLFYLTDCGLTDTLKSADLVIDFFEHGSEA 61
Db 3 DKTVPVFSMAELRDGSRQDEFFREWAR-RGVFYLTVGATERDHRVATDTAMOFFFAQGTAE 61
QY 62 EKRAVTSVPVPTMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLFPSPGDFGRITWQ 121
Db 62 EKQAVTTKVPVPTMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLFPSPGDFGRITWQ 121
QY 122 YFDRQYTSRAVARAVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRMA 181
Db 122 YFDSLRYRAAQETARLVLTAAAGTYDGEDLTLLDCDPVLRVFPPEHRAAAEYEPERRMA 181
QY 182 PHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVTTGGQVAP 241
Db 182 PHYDLSIIITFIHQTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVTTGGQVAP 241
QY 242 RHVVAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGN 301
Db 242 NHVWSPDASMLKGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGN 301
QY 302 YVNI 305
Db 302 YVTM 305

RESULT 8

US-09-791-537-10913
; Sequence 10913, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10913
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Streptomyces lactamdurans
US-09-791-537-10913

Query Match 59.0%; Score 965.5; DB 22; Length 310;
Best Local Similarity 61.2%; Pred. No. 2e-101;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAELOQGLHODEFRRCRLDKGLFYLTDCGLTDTLKSADLVIDFFEHGSEA 61
Db 3 DKTVPVFSMAELRDGSRQDEFFREWAR-RGVFYLTVGATERDHRVATDTAMOFFFAQGTAE 61
QY 62 EKRAVTSVPVPTMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLFPSPGDFGRITWQ 121
Db 62 EKQAVTTKVPVPTMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLFPSPGDFGRITWQ 121
QY 122 YFDRQYTSRAVARAVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRMA 181
Db 122 YFDSLRYRAAQETARLVLTAAAGTYDGEDLTLLDCDPVLRVFPPEHRAAAEYEPERRMA 181
QY 182 PHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVTTGGQVAP 241
Db 182 PHYDLSIIITFIHQTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVTTGGQVAP 241
QY 242 RHVVAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGN 301
Db 242 NHVWSPDASMLKGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGN 301
QY 302 YVNI 305
Db 302 YVTM 305

RESULT 9
US-09-791-537-10914
; Sequence 10914, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 10914
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-791-537-10914

Query Match 57.4%; Score 939.5; DB 22; Length 318;
Best Local Similarity 58.6%; Pred. No. 2.1e-98;
Matches 180; Conservative 43; Mismatches 81; Indels 3; Gaps 1;

QY 2 DTTVPTFSLAELOOGLHDEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVIDFFEHGSEA 61
DB 3 DTPVPFNLAAALREGADQEKRECVTGMGVFYLTYGAGDKDHLATDTAMDFFANGTEA 62
QY 62 EKRAVTSVPVMTWRRGFTGLESESTAIQTNTGSYSDYSCYSGMTADNLFPSPGDFGRITQ 121
DB 63 EKAATVDTPTWRRGYSALAEASTAQVTRTGSYDYSMSFSGISGNVFPSPFERVWTE 122
QY 122 YPDROYTASRAVAREVLRTGTEPDGVEAFDCEPLLRFRYPVPOVPEHRSABEOLP 178
DB 123 YFDKLVAAQETARLVLTAGGGYDAEILVGLDELLADPVLRLRYFPVPEHRSABEHP 182
QY 179 RMAPHYDLSMVTLIQOTPCANGFVSQAQVGGAFDLPYRPDAVLVFCGAIAITLVGGQV 238
DB 183 RMAPHYDLSIITFIHQTPCANGFVSQAQVGGAFDLPYRPDAVLVFCGAIAITLVGGQV 242
QY 239 KAPRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGDFVSLDGETATQDWI 298
DB 243 PAPRHVRSFGAGWRGSDRTSSVFFLRPNADFTFSVPLARECGDFVSLDGETATQDWI 302
QY 299 GGNVNI 305
DB 303 GTNVITM 309

RESULT 10
US-09-791-537-4795
; Sequence 4795, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4795
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Lysobacter lactamgenus
US-09-791-537-4795

Query Match 57.2%; Score 937; DB 22; Length 319;
Best Local Similarity 58.0%; Pred. No. 4.1e-98;
Matches 177; Conservative 45; Mismatches 83; Indels 0; Gaps 0;

QY 2 DTTVPTFSLAELOOGLHDEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVIDFFEHGSEA 61
DB 3 DSGIQIFDLDELEHGVRLDSPKSLFERGVFVYREDDSIKTEHAKAMDAVMDLFENGSAE 62
QY 62 EKRAVTSVPVMTWRRGFTGLESESTAIQTNTGSYSDYSCYSGMTADNLFPSPGDFGRITQ 121
DB 63 QKNALRNLPNRRGFSDLAEASTAITKGGYEYTSWYISGLTDLNLPSPFAFEAITWG 122
QY 122 YPDROYTASRAVAREVLRTGTEPDGVEAFDCEPLLRFRYPVPOVPEHRSABEOLP 181
DB 123 YPDROYTASRAVAREVLRTGTEPDGVEAFDCEPLLRFRYPVPOVPEHRSABEOLP 182
QY 182 RMAPHYDLSMVTLIQOTPCANGFVSQAQVGGAFDLPYRPDAVLVFCGAIAITLVGGQV 241
DB 183 RMAPHYDLSIITFIHQTPCANGFVSQAQVGGAFDLPYRPDAVLVFCGAIAITLVGGQV 242
QY 242 RHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGDFVSLDGETATQDWI 301
DB 243 RHQVASPSMQORVGSRTSSVFFLRPNADFTFSVPLARECGDFVSLDGETATQDWI 302
QY 302 YVNI 306
DB 303 YVNI 307

RESULT 11
US-09-791-537-94436
; Sequence 94436, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 94436
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Acremonium chrysogenum
US-09-791-537-94436

Query Match 56.6%; Score 926.5; DB 22; Length 332;
Best Local Similarity 56.7%; Pred. No. 7.1e-97;
Matches 177; Conservative 43; Mismatches 91; Indels 1; Gaps 1;

QY 1 MDTVPTFSLAELOOGLHDEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVIDFFEHGSE 60
DB 1 MTSKVPVFRLLDLSKSKVLTAEAVTTKGIYLTESGLVDDDDHTSARETCVDFKNGSE 60
QY 61 EKRAVTSVPVMTWRRGFTGLESESTAIQTNTGSYSDYSCYSGMTADNLFPSPGDFGRITQ 120
DB 61 EKRAVTLADRNRARGFSALAEASTAVTETGKYSYSCYSGMIGNLPNRRGDFVWQ 120
QY 121 YPDROYTASRAVAREVLRTGTEPDG-GVEAFDCEPLLRFRYPVPOVPEHRSABEOLP 179
DB 121 YPDROYTASRAVAREVLRTGTEPDG-GVEAFDCEPLLRFRYPVPOVPEHRSABEOLP 180
QY 180 RMAPHYDLSMVTLIQOTPCANGFVSQAQVGGAFDLPYRPDAVLVFCGAIAITLVGGQV 239
DB 181 RMAPHYDLSIITFIHQTPCANGFVSQAQVGGAFDLPYRPDAVLVFCGAIAITLVGGQV 240
QY 240 APRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGDFVSLDGETATQDWI 299
DB 241 APRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGDFVSLDGETATQDWI 300
QY 300 GNVNIRRTSKA 311
DB 301 GNVNIRRTSKA 312

```

RESULT 12
US-09-791-537-4796
; Sequence 4796, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4796
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Lysobacter lactamgenus
US-09-791-537-4796

```

[illegible]

```

RESULT 13
US-09-791-537-89857
; Sequence 89857, Application us/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 89857
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Streptomyces jumonjinensis
US-09-791-537-89857

```

Query Match	10.9%;	Score 179;	DB 22;	Length 329;
Best Local Similarity	24.9%;	Pred. No. 9.7e-11;		

	Matches	72; Conservative	45; Mismatches	130; Indels	42; Gaps	11;
Qy	5	VPTSL-----AEIQQLHQHDEPRCLDKGLFYLTDCGLTDTLTKLSAKDLVIDFFEH	57			
Db	10	VPTIDISPLSGDDAKAQKRVQAQ-EINKAARGSGFFYASNHGV---DVQLLDVVVNEPHN	65			
Qy	58	GSAEK-----RAVTSVPVTMRGF----TGLESESTAOITNTGYSYDSMCSYMGTD-	107			
Db	66	MSDOEKHDLAINAYKNDNPHVRNGYYKAIGKKAVESFCVLNPSFSDDHFMWIKSETPMHE	125			
Qy	108	NLFPSGD----FGRIWTQYFDQYTASRAVARREVLRATETPDGGVEAFDCE-----P	157			
Db	126	VNLWPDEEKHPFRPFCCEDYYRQLRLSLTVIMRGYALALGRERDFPEALAAEADTLSSVS	185			
Qy	158	LLEFRYPQPVEHRS-AEEOPLEMAHYDLMSWTLLOOTECANGFVFSLOAEVGGAFTDLP	216			
Db	186	LIRYPLEEYPPVKTAGDGTKLSFEDHLDMVMTLVLYQTENVQVN----LQVETVDWGQDIP	241			
Qy	217	YRPDAVLVFCAIATLVTGGQVKAPRHVAAPRRDQIAGSRSTSVPFL	265			
Db	242	RSEDDELNVCGTYMGHTTHDYFPAPHRV-----KFINAERISLPFFL	284			

```

RESULT 14
US-09-791-537-9085
; Sequence 9085, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blomonic, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL
; TITLE OF INVENTION: METHODS OF USE T
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9085
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Streptomyces griseus
US-09-791-537-9085

```

	Query Match	10.8%	Score 176;	DB 22;	Length 329;
	Best Local Similarity	24.5%;	Pred. No. 2.2e-10;		
	Matches	73;	Conservative	39;	Mismatches 126; Indels 60; Gaps 12;
Qy	5 VPTFSLAEQLQGHD-----BFRRCLRDKGLFYLTDCGLTDTTELKSADKLVIDFFEHG	58			
Dd	10 VPTIDISPLSGGDADDKKRVAQEINKACRESSGFYASHHGI--DVQLLKDVMVNEFHRTM	66			
Qy	59 SEAEK-----RAVTSPVPTMRGTFTGLESESTAQITNTGSYSDSMCXSMGTADNLFP-	111			
Dd	67 TDEKYVDLAINAYKNPNTRNGY-----YMVKGGKAVESWCYLNPFSSEDHPQI	117			
Qy	112 -SG---DFGRIW-----TOYDRQYTASRAVARVLRTATGPDPGVGEA---	151			
Dd	118 RSGTPMEHGNIPDEKRQHQRFRPCEQYRYDRVFSLSKVLMRGFALALG-KPEDPFDA	176			
Qy	152 ---FLDCEPLLRFYPQPVEHSABE-QPLRMAPHYDLMSWTLIIQTPCANGFVSQAQ	207			
Dd	177 LADTLSANTLIHYPLEDYPPVKTPDGTKLSFEDHLDDVSMIVLPFTEVN-----IQVE	232			
Qy	208 VGGAFTDLPYRPDAVLVFCGAIATLVTTGGOVKAPRRHHVAAARRDQIAGSSRTSSVFLL	265			
Dd	233 TADGWODLPTSGENFLVCCTYMGYLINDYFPAPNHRV-----KFINAERLSLPPFL	284			

RESULT 15
US-10-369-493-12239
; Sequence 12239, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12239
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-12239

Query Match 10.3%; Score 168; DB 29; Length 366;
Best Local Similarity 24.7%; Pred. No. 2.1e-09;
Matches 85; Conservative 41; Mismatches 158; Indels 60; Gaps 14;
QY 1 MDTTTFSLAELOQGLHQ-----DFERRCLRDKGLFYLTDGCLTDELKSAKDLVIDEF 55
DB 28 MPRIVPVLDSRLSEQASERRTFADLRASARDIGFFYLAGHGISWAEISEVLTAQRQFF 87
QY 56 EHGSEAEKRAVTSVPVPMRSGFTGLESESTAQITNTGYSYSDYSCYMGTDNLFPSPGDF 115
DB 88 AL-PEADKLAIEVMKSSQFRGYTRAGELTK-----GREDWREQLDIGVERQAIAGPG 140
QY 116 GRIWTO-YFDRQYTAS-----RAVAREVLRATCTEPDGGVEAFDCEPLL 159
DB 141 TPATRLQGNQWPAALPDLKPALLAWQSKVTAVAIRLLKAFQAQSLDQPEDAF---DPI- 196
QY 160 RFRYFPQVPEHR-SAEQPLR-----WAPHYDLSWVTLIOOTPCANGFVSLQAEVGG 210
DB 197 -----YSSEPNHRMKIVRYPGRDTTGGDQGVGAHKDGGFLTLLOQ-----DDNKGQVYDYG 248
QY 211 AFTDLPVRPDVLFVFCGAIATLVTGQGVKAPRHHVAAPRRDQIAGSSRTSSVFFLRPNAD 270
DB 249 SWVDVDPFGTLVNVNIGELLELASNGYLRTVHRVTPP-----AGVERISVPEFFSARLD 304
QY 271 FT-----FSVPLARECGFDVSLDGETATFODMIGGNVNIIRTS 309
DB 305 ATIPLLGLSELAQAARGPAS-DPDNPLFRD-VGTNVLKSRIRS 346

Search completed: February 3, 2004, 17:41:40
Job time : 176 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:37:44 ; Search time 22 Seconds
(without alignments)
1088.984 Million cell updates/sec

Title: US-09-582-486-1

Perfect score: 1637

Sequence: 1 MDTVTFTSLAEQLQGLHQD.....ATFQDWIGGNYNVRTSKA 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374145 seqs, 77034267 residues

Total number of hits satisfying chosen parameters: 374145

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New.*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pap.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pap.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pap.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pap.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pap.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pap.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	174.5	10.7	345	6	US-10-679-063-16188
2	169	10.3	412	6	US-10-679-063-16144
3	160	9.8	342	6	US-10-679-063-22244
4	155	9.5	375	6	US-10-425-114A-58493
5	154	9.4	352	6	US-10-425-114A-63175
6	153	9.3	371	6	US-10-425-114A-62424
7	153	9.3	372	6	US-10-425-114A-53454
8	153	9.3	372	6	US-10-425-114A-72059
9	152	9.3	379	6	US-10-679-063-13321
10	148.5	9.1	389	6	US-10-481-381-4
11	137.5	8.4	335	6	US-10-679-063-9242
12	137.5	8.4	380	1	PCT-US03-24364-36
13	136.5	8.3	380	6	US-10-257-494A-16
14	135.5	8.3	405	6	US-10-739-930-7737
15	133	8.1	385	6	US-10-425-114A-66945
16	131.5	8.0	371	7	US-60-478-196-3232
17	131	8.0	356	6	US-10-425-114A-64811
18	131	8.0	356	6	US-10-739-930-7228
19	131	8.0	378	6	US-10-739-930-6723
20	131	8.0	378	6	US-10-257-494A-15
21	131	8.0	542	5	US-09-614-150A-11103
22	130.5	8.0	353	6	US-10-425-114A-61294
23	130.5	8.0	382	6	US-10-425-114A-66617
24	128.5	7.8	342	6	US-10-679-063-8477
25	128	7.8	346	6	US-10-679-063-15920
26	127	7.8	334	6	US-10-679-063-9241

ALIGNMENTS

RESULT 1

US-10-679-063-16188
; Sequence 16188, Application US/10679063
; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; PRIOR FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 16188
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Rhodospirillum rubrum
US-10-679-063-16188

Query Match 10.7%; Score 174.5; DB 6; Length 345;
Best Local Similarity 26.8%; Pred. No. 8e-09;
Matches 81; Conservative 32; Mismatches 134; Indels 55; Gaps 12;

Qy	6	PTFSL-AELQOGLHODEF-----RRLCDKGLVLTDCGLTDTLTKSAKOLVIDPFEHSE	60
Db	13	PVLSLKAGEGEAARRDEFLATLRQAARDPCAFYLEGHGIDPSVIDQVEALSRRFPAL-PE	71
Qy	61	AEKRAVTSVPVTMRRTGFTGLESESTAQITNTGSYSYSCYSMGTADNLPSPGDFGRWT	120
Db	72	DEKRAIDMVNSPHFHGYTRVGAELTR-----GAPDREQLDIGSERPLLPGQPDTPAWA	125
Qy	121	QY-----FDRQVTA-SRAVAREVLRTGTEPDGGEAFIDC-----	155
Db	126	RLQGNQWPAALPDLRAAIVRLQAEALTAVALLERIALALGERAD----FFADLYEGGP	181
Qy	156	EPLLRFYEPQVPEHSEAEQPLRMAPHVDLSMVTLLQTPPCANGFVSLQAEVGGFTDL	215
Db	182	DQLKTIIRYP-----GGAAGEGDGQGVGHKDSGLLTFLVLRDRG-----GLQIERE	233
Qy	216	PYRPDAVLCGAIATLVGTGGQVKAPRHVAAPRRQIAGSSRTSSVFLRPNADFTFSV	275
Db	234	PRPGTFVANNIGELLELATNGYLKATVHRVVSPPAD-----SDRLSTAFPL--CARL	287
Qy	276	PL 277	
Db	288	PL 289	

RESULT 2

US-10-679-063-16144

APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaka, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114A
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 63175
LENGTH: 352
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3062-030-F8_FLI.pep
US-10-425-114A-63175

Query Match 9.4%; Score 154; DB 6; Length 352;
Best Local Similarity 23.4%; Pred. No. 7.7e-07;
Matches 78; Conservative 49; Mismatches 119; Indels 88; Gaps 16;
QY 4 TVPTFSLAELOQGLHODEFRCL-----RDKGLFYLTDCGLTDTTELKSAKDLVIDDFEHGS 59
DB 41 TSATVSLPIVDLSLGRDEVRRAILEAGKEIGFQVNVHGVSLQAEQMETVCOEPR--L 98
QY 60 EAEKRAVTSVPVPTMRGFTGLESESTAQIT-----NTGSYSDYSCYSGMTADNLPP 111
DB 99 PAEDKA-----GLYSEDTRATRIYSSMTFDTGGEKYWRDCLRLACS---PP 142
QY 112 S-GDFGRIW-----TQYFDRQYTRASRAVAREVLR-----ATGTEPD-----GGVEAF 152
DB 143 AVGDSAAAMPDKPRRLREVVERFTVQTRGLGMEILRLCEGLGLRDPYLEGDISG-- 198
QY 153 LDCEPLLRFRYPQVPEHRSABEQPLRMAPHYDLSMTLIQOQPCANGFVSLQAEVGGAF 212
DB 199 ---DVLVHNHYPPCPDPNAT---LGLPPHCDRNLLTL-----LPSMWPG-- 238
QY 213 TDLPYR-----PDAVLVFCGAIATLVGGQVKAPRRHVAAPRRDQIAGSSRTSSVF 263
DB 239 LEVAYRGDWIRVPEVPGAFVNVFGCQLEVVNTGILKSIHRVTN-----LGVARTTVAT 293
QY 264 FLRPNADFTFSVPLARECGFDVSLDGETATFQDW 297
DB 294 FIMPTTDCILIG-PAABFLSDNDPPCYRTLTFGDF 326

RESULT 6
US-10-425-114A-62424
; Sequence 62424, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaka, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62424
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3357-058-F2_FLI.pep
US-10-425-114A-62424

Query Match 9.3%; Score 153; DB 6; Length 371;

Best Local Similarity 23.4%; Pred. No. 1e-06;
Matches 78; Conservative 48; Mismatches 120; Indels 88; Gaps 16;
QY 4 TVPTFSLAELOQGLHODEFRCL-----RDKGLFYLTDCGLTDTTELKSAKDLVIDDFEHGS 59
DB 60 TSATVSLPIVDLSLGRDEVRRAILEAGKEIGFQVNVHGVSLQAEQMETVCOEPR--L 117
QY 60 EAEKRAVTSVPVPTMRGFTGLESESTAQIT-----NTGSYSDYSCYSGMTADNLPP 111
DB 118 PAEDKA-----GLYSEDTRATRIYSSMTFDTGGEKYWRDCLRLACS---PP 161
QY 112 S-GDFGRIW-----TQYFDRQYTRASRAVAREVLR-----ATGTEPD-----GGVEAF 152
DB 162 AVGDSAAAMPDKPRRLREVVERFTVQTRGLGMEILRLCEGLGLRDPYLEGDISG-- 217
QY 153 LDCEPLLRFRYPQVPEHRSABEQPLRMAPHYDLSMTLIQOQPCANGFVSLQAEVGGAF 212
DB 218 ---DVLVHNHYPPCPDPNAT---LGLPPHCDRNLLTL-----LPSMWPG-- 257
QY 213 TDLPYR-----PDAVLVFCGAIATLVGGQVKAPRRHVAAPRRDQIAGSSRTSSVF 263
DB 258 LEVAYRGDWIRVPEVPGAFVNVFGCQLEVVNTGILKSIHRVTN-----LGVARTTVAT 312
QY 264 FLRPNADFTFSVPLARECGFDVSLDGETATFQDW 297
DB 313 FIMPTTDCILIG-PAABFLSDNDPPCYRTLTFGDF 345

Query Match 9.3%; Score 153; DB 6; Length 372;
Best Local Similarity 23.4%; Pred. No. 1e-06;
Matches 78; Conservative 48; Mismatches 120; Indels 88; Gaps 16;
QY 4 TVPTFSLAELOQGLHODEFRCL-----RDKGLFYLTDCGLTDTTELKSAKDLVIDDFEHGS 59
DB 61 TSATVSLPIVDLSLGRDEVRRAILEAGKEIGFQVNVHGVSLQAEQMETVCOEPR--L 118
QY 60 EAEKRAVTSVPVPTMRGFTGLESESTAQIT-----NTGSYSDYSCYSGMTADNLPP 111
DB 119 PAEDKA-----GLYSEDTRATRIYSSMTFDTGGEKYWRDCLRLACS---PP 162
QY 112 S-GDFGRIW-----TQYFDRQYTRASRAVAREVLR-----ATGTEPD-----GGVEAF 152
DB 163 AVGDSAAAMPDKPRRLREVVERFTVQTRGLGMEILRLCEGLGLRDPYLEGDISG-- 218
QY 153 LDCEPLLRFRYPQVPEHRSABEQPLRMAPHYDLSMTLIQOQPCANGFVSLQAEVGGAF 212
DB 219 ---DVLVHNHYPPCPDPNAT---LGLPPHCDRNLLTL-----LPSMWPG-- 258
QY 213 TDLPYR-----PDAVLVFCGAIATLVGGQVKAPRRHVAAPRRDQIAGSSRTSSVF 263

```
Db 259 LEVAYGDMIRVEPVGAFVNFVGCQLEVTNGILKSIEHRVTN-----LGVARTTAT 313
QY 264 FLRPNADFTSVPLARECGFDVSLDGETATFQW 297
Db 314 FIMPTDCLIG-PAABFLSDNDPPCYRTLTFGDF 346

RESULT 8
US-10-425-114A-72059
; Sequence 72059, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 72059
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3245-358-A7_FLI.pep
US-10-425-114A-72059
```

```
Query Match 9.3%; Score 153; DB 6; Length 372;
Best Local Similarity 23.4%; Pred. No. 1e-06;
Matches 78; Conservative 48; Mismatches 120; Indels 88; Gaps 16;

QY 4 TVTFSLAELQOGLHODEFRCL-----ROKGLFVLTDCGLTDTTELKSAKDLVIDDFEHGS 59
Db 61 TSATVSLPIVDLSDGRDEVRRALEAGKEIGFFQVNVHGVSLAMQDMETVCOEFPF--L 118
QY 60 EAKRAVTSVPVMTMRGFTGLESESTAQIT-----NTGSYSYSDYSCVSMGTADNLF 111
Db 119 PAEDKA-----GLYSEDIGTRATRIYSTMTDGTGGEKYWRDCLRLACS---FP 162
QY 112 S-GDGRIM-----TOYFDROYTASRAVAREVLR-----ATGTEPD-----GGVEAF 152
Db 163 AVGDSAAAWPKPRRLREVVERFTVQTRGLGMBILRLCEGLGLRDPYLEGDISGG--- 218
QY 153 LDCEPLLRFRYPQVPEHRSABEQPLRMAPHYDLSMVTLLIQOTPCANGFVSLQAEVGGAF 212
Db 219 ---DVLVHNHYPCDPDNAT-----LGLPPHCDRNLLTL-----LPSWVPG-- 258
QY 213 TDLFPR-----PDAVLVFCGAIATLVGGQVKAPRHVHAAPRRDQTAGSRTSSVF 263
Db 259 LEVAYGDMIRVEPVGAFVNFVGCQLEVTNGILKSIEHRVTN-----LGVARTTAT 313
QY 264 FLRPNADFTSVPLARECGFDVSLDGETATFQW 297
Db 314 FIMPTDCLIG-PAABFLSDNDPPCYRTLTFGDF 346
```

```
RESULT 9
US-10-679-063-13321
; Sequence 13321, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679, 063
; CURRENT FILING DATE: 2003-10-02
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
```

```
; SEQ ID NO 13321
; LENGTH: 379
; TYPE: PRT
US-10-679-063-13321

Query Match 9.3%; Score 152; DB 6; Length 379;
Best Local Similarity 24.3%; Pred. No. 1.3e-06;
Matches 67; Conservative 40; Mismatches 105; Indels 64; Gaps 12;

QY 30 GLFYLTDCGLTDTTELKSAKDLVIDDFEHGSEAEKRAVTSVPVMTMR-----GFTG- 79
Db 92 GFFLVNHGVDDKLIHAHAHQYIDYFFELPMSAQRA-----QRKVGSHCGYASSFTGR 144
QY 80 -----LESESTAQITNTGSYSYSDYSCVSMGTADNLFPSGDFGRITWTOYFDBQYTA 129
Db 145 FSKFLPWKETLSFRSSAOPDSSNIQDY--LCNTMG--EDFKP---FGKYVQYCDAMSTL 198
QY 130 SRAVAREVLRATGTEPDGGVEAFDCEPLLRFRYPF--QVPEHRSABEQPLRMAPHYDLS 187
Db 199 SLGIMELLGMSLVGSQGHYRFFENESIMRLNYYPCKCP-----DLTLGTGPHCDPT 252
QY 188 MVTLLIQOTPCANGFVSLQAEVGG--AFTDLPYR-----DAVLVFCGAIATLVGGQVKA 240
Db 253 SLTILHQD-----QVGGLOQVFDDEWRSITPNENAFVNIQDTFMALSNGRYKS 301
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVP 276
Db 302 CLHRAVWNSK-----TPRKSALFFLCRKNDKVVSPP 332
```

```
RESULT 10
US-10-481-381-4
; Sequence 4, Application US/10481381
; GENERAL INFORMATION:
; APPLICANT: Okawa, Miho
; APPLICANT: Matsuo, Makoto
; APPLICANT: Ashikari, Motoyuki
; TITLE OF INVENTION: sd1 gene involved in plant semidwarfing and uses thereof
; FILE REFERENCE: SHZ-018US
; CURRENT APPLICATION NUMBER: US/10/481,381
; CURRENT FILING DATE: 2003-12-18
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-481-381-4
```

```
Query Match 9.1%; Score 148.5; DB 6; Length 389;
Best Local Similarity 24.8%; Pred. No. 2.9e-06;
Matches 64; Conservative 40; Mismatches 113; Indels 41; Gaps 10;

QY 30 GLFYLTDCGLTDTTELKSAKDLVIDDFEHGSEAEKRAVTSVPVMTMRGFTGLESE----STA 86
Db 94 GFFQVSEHGVDAALARAALDGSDFRLPLAEKRR--RRVEGTVSQYTSAAHADRFAASKL 151
QY 87 QITNTGSYSYSDYSCVSMGTAD---NLPPSGDF---GRIWTOYFDROYTASRAVAREVLR 139
Db 152 PKWETLSFGFHDRAAAPVVDYFSSTLGP--DFAPMGVYQYKCEMKELSLTIMELLEL 209
QY 140 ATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSABEQPLRMAPHYDLSMVTLLIQOTPCAN 199
Db 210 SLGVEGYVREFFADSSSIMRCNYYPCCPE---PERTLGTGPHCDPTALTIL----- 258
QY 200 GFVSLQAEVGG--AFTDLPYR-----PDAVLVFCGAIATLVGGQVKAPRHVHAAPRRDQ 252
Db 259 ----LQDDVGGLEVLVDGEWRPVPVPGAMVINIGDTFMALSNGRYKSLHRAVAVNQRE 314
QY 253 IAGSSRTSSVFFLRPNAD 270
```

```
Db 315 -----RRSLAFLCPRED 327
| | | | |
RESULT 11
US-10-679-063-9242
; Sequence 9242, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 9242
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Citrus unshiu
US-10-679-063-9242
Query Match 8.4%; Score 137.5; DB 6; Length 335;
Best Local Similarity 21.8%; Pred. No. 2.8e-05;
Matches 63; Conservative 47; Mismatches 124; Indels 55; Gaps 10;
QY 5 VPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKLSAKDLVIDFFEHGSAEKR 64
| | | | |
Db 43 IPTIDDDPQDRLVRSIAASREWGIFQVTHNGIFSDLI CKLQAVGKEFFELPQB-EKE 101
| | | | |
QY 65 AVTSPVPTMRGFTGLESESTAQITNTGYSYDSCMYSMGTADNLPFGSDGFRWT- 120
| | | | |
Db 102 VYSRPADA--KDVQGYGTLQKEVGKSKWVDHL-----FHRVWPPSSI 143
| | | | |
QY 121 --QYFDRQYTSARAVAREVUR-----ATGTEPDGCV--EAF--LDCBPLLR 161
| | | | |
Db 144 NYRFPKPNPSPSYRANVEYAKYMRVVDKLFYTLGLGVEGGVLRKEAAGGDDIEYMLKI 203
| | | | |
QY 162 RYFPQVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDTLPYRPDA 221
| | | | |
Db 204 NYTPPCR-----PDLAGVVAHTDLSALTVL-----VPEVPGLOVFKDDRWDIAKIPNA 255
| | | | |
QY 222 VLVFCGAIAITLVGGQVKAPRHVAAPRRDQIAGSSRTSSVFFLRPNAD 270
| | | | |
Db 256 LVTHIGDQIBILSNGKYKAVLHRTTVNK-----DKTRMSWPVLEPPAD 299
| | | | |
RESULT 12
PCT-US03-24364-36
; Sequence 36, Application PC/TUS0324364
; GENERAL INFORMATION:
; APPLICANT: BASF PLANT SCIENCE GMBH
; TITLE OF INVENTION: SUGAR AND LIPID METABOLISM REGULATORS IN PLANTS IV
; FILE REFERENCE: 16313-0236
; CURRENT APPLICATION NUMBER: PCT/US03/24364
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US 60/400,803
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
PCT-US03-24364-36
Query Match 8.4%; Score 137.5; DB 1; Length 380;
Best Local Similarity 23.4%; Pred. No. 3.3e-05;
Matches 69; Conservative 38; Mismatches 143; Indels 45; Gaps 11;
QY 5 VPTFSLAELOOQ---LHQDEFR---RCLRDKGLFYLTDCGLTDTLTKLSAKDLVIDPFEHG 58
| | | | |
Db 58 VPLIDLAGFLSGDCLASEATRLVSKAATKHGFFLITNHGI-DESLLSRAYLHMDSPFKA 116
| | | | |
QY 59 SRAEKAVTSPVPTMR-----GFTG-LESESTAQITNTGYSYDSCMYSMGTADN 108
| | | | |
Db 117 PACEKQ-----KAQRKMGESSGYASSFVGRFSSKLPWKETLSFKFSPBEKIHSTQTKVD- 169
| | | | |
QY 109 LFPSS-----GDFGRITWQYFDRQYTSARAVAREVLRATGTDPDGGVEAFDCEPLLR 161
| | | | |
Db 170 -FVSKMGDGYEDFGKYQYEAEMNTLSLKIMELLGMSLGVERRYFKFEFFEDSDSIFRL 228
| | | | |
QY 162 RYFPQVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDTLPYRPDA 221
| | | | |
Db 229 NYTPQCKQ-----PELAGTGPHCDPSTLSILHODQVG-----GLQVFDNKKWQSIIPNPHA 280
| | | | |
QY 222 VLVFCGAIAITLVGGQVKAPRHVAAPRRDQIAGSSRTSSVFFLRPNADTFFSVP 276
| | | | |
Db 281 FVNIIGDTFMALTNGRYKSCCLHRAVNVNRE-----RKTFAFFLCPKGEKVVKPP 330
| | | | |
```

```
QY 59 SRAEKAVTSPVPTMR-----GFTG-LESESTAQITNTGYSYDSCMYSMGTADN 108
| | | | |
Db 117 PACEKQ-----KAQRKMGESSGYASSFVGRFSSKLPWKETLSFKFSPBEKIHSTQTKVD- 169
| | | | |
QY 109 LFPSS-----GDFGRITWQYFDRQYTSARAVAREVLRATGTDPDGGVEAFDCEPLLR 161
| | | | |
Db 170 -FVSKMGDGYEDFGKYQYEAEMNTLSLKIMELLGMSLGVERRYFKFEFFEDSDSIFRL 228
| | | | |
QY 162 RYFPQVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDTLPYRPDA 221
| | | | |
Db 229 NYTPQCKQ-----PELAGTGPHCDPSTLSILHODQVG-----GLQVFDNKKWQSIIPNPHA 280
| | | | |
QY 222 VLVFCGAIAITLVGGQVKAPRHVAAPRRDQIAGSSRTSSVFFLRPNADTFFSVP 276
| | | | |
Db 281 FVNIIGDTFMALTNGRYKSCCLHRAVNVNRE-----RKTFAFFLCPKGEKVVKPP 330
| | | | |
RESULT 13
US-10-257-494A-16
; Sequence 16, Application US/10257494A
; GENERAL INFORMATION:
; APPLICANT: AN, GYN-HEUNG
; APPLICANT: KANG, HOUG-GYU
; APPLICANT: JUN, SUNG-HOON
; APPLICANT: KIM, JUNYUL
; APPLICANT: JUNG, KI-HWAN
; APPLICANT: PARK, YONG-JOO
; APPLICANT: LEE, SANG-YEB
; TITLE OF INVENTION: A PROMOTER CV200XP WHICH REGULATES THE INTEGRMENT-SPECIFIC
; TITLE OF INVENTION: EXPRESSION OF A GIBBERELLIN 20-OXIDASE GENE IN DEVELOPING
; TITLE OF INVENTION: SEEDS OF WATERMELON AND A METHOD FOR GENERATING SEEDLESS
; TITLE OF INVENTION: FRUITS USING THE PROMOTER
; FILE REFERENCE: 7022-0003
; CURRENT APPLICATION NUMBER: US/10/257,494A
; CURRENT FILING DATE: 2003-08-16
; PRIOR APPLICATION NUMBER: PCT/KR00/01127
; PRIOR FILING DATE: 2000-10-09
; PRIOR APPLICATION NUMBER: KR 2000/18483
; PRIOR FILING DATE: 2000-04-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-257-494A-16
Query Match 8.3%; Score 136.5; DB 6; Length 380;
Best Local Similarity 23.4%; Pred. No. 4.1e-05;
Matches 69; Conservative 38; Mismatches 143; Indels 45; Gaps 11;
QY 5 VPTFSLAELOOQ---LHQDEFR---RCLRDKGLFYLTDCGLTDTLTKLSAKDLVIDPFEHG 58
| | | | |
Db 58 VPLIDLAGFLSGDCLASEATRLVSKAATKHGFFLITNHGV-DESLLSRAYLHMDSPFKA 116
| | | | |
QY 59 SRAEKAVTSPVPTMR-----GFTG-LESESTAQITNTGYSYDSCMYSMGTADN 108
| | | | |
Db 117 PACEKQ-----KAQRKMGESSGYASSFVGRFSSKLPWKETLSFKFSPBEKIHSTQTKVD- 169
| | | | |
QY 109 LFPSS-----GDFGRITWQYFDRQYTSARAVAREVLRATGTDPDGGVEAFDCEPLLR 161
| | | | |
Db 170 -FVSKMGDGYEDFGKYQYEAEMNTLSLKIMELLGMSLGVERRYFKFEFFEDSDSIFRL 228
| | | | |
QY 162 RYFPQVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDTLPYRPDA 221
| | | | |
Db 229 NYTPQCKQ-----PELAGTGPHCDPSTLSILHODQVG-----GLQVFDNKKWQSIIPNPHA 280
| | | | |
QY 222 VLVFCGAIAITLVGGQVKAPRHVAAPRRDQIAGSSRTSSVFFLRPNADTFFSVP 276
| | | | |
Db 281 FVNIIGDTFMALTNGRYKSCCLHRAVNVNRE-----RKTFAFFLCPKGEKVVKPP 330
| | | | |
```

RESULT 14
US-10-739-930-7737
; Sequence 7737, Application US/10739930
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 7737
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(405)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C24992_1.p
US-10-739-930-7737

Query Match 8.3%; Score 135.5; DB 6; Length 405;
Best Local Similarity 24.2%; Pred. No. 5.5e-05;
Matches 61; Conservative 41; Mismatches 123; Indels 27; Gaps 8;
QY 30 GLFYLTDGLTDLTSLAKDLVDFPHGSEAEKRAVTSVPVPMRGFTGLESESTA--- 86
Db 113 GPFQVGHGVDAAALGRAALDGSDFRLPLAEKQRA--RRVPGTVSGYTSAHADRFPAKL 170
QY 87 --QITWTGYSYDYS-----MCYSMTADNLPFSGDFGRITQYDFDRQYTASRAVAREVLR 139
Db 171 PKETLSFGHDGAXXPVVVDYVGLGQDFE--PMGWYQRYCEEMKELSLTIMELLEL 228
QY 140 ATCTEPDGGV-EAFLDCEPLRLFRYPQVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCA 198
Db 229 SLGVELRGYYREFEFDSRSMRCNYPPCPPE---PERTLGTGPHCDPTALTILLQDXDV 284
QY 199 NGFVSLOAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKAPRHHVAAPRRDQIAGSSR 258
Db 285 GG---LEVLDVGWRPVPVPGAMVINIGDTFMALSNGRYKSLHRAVNVQR-----RAR 336
QY 259 TSSVFFLRPNAD 270
Db 337 RSLAFFLCRPED 348

RESULT 15
US-10-425-114A-66945
; Sequence 66945, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 66945
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4756-039-A6_FLI.pgp
US-10-425-114A-66945

Query Match 8.1%; Score 133; DB 6; Length 385;

Best Local Similarity 21.9%; Pred. No. 9e-05;
Matches 61; Conservative 43; Mismatches 140; Indels 34; Gaps 9;
QY 5 VPTFSLAEIQQ---GLHODEPRRCRLRDKGLFYLTCGLTDTLTKSAKDILVIDFFEHGSEA 61
Db 67 IPVIDVTELQHDHVDGL--DKLRLACEEWGFFOVVNHGIAHELLEMERLAREFFMLPLEE 124
QY 62 EKRAVTSVPVPMRGFTGLESESTAQITNTGYSYDYSYCMYSGMT-----ADNLFPSS--G 113
Db 125 KEKYPMAP-----GGIQGYGHAFVSEDDQKLDWCNNMLALGVPEPASIRQPLWPTAPA 176
QY 114 DFGRIWTQYDFDRQYTASRAVAREVLRATGTEPDGGVEAFDCEPLRLFRYPQVPEHRS 173
Db 177 GFGETLEAYSAEVGBELCRRLLARIAETTLGLAPATTADMFGAEAVQVRMNFYPPCPR--- 232
QY 174 EEQPLRMAPHYDLSMVTLIQOTPCANGFVSLQAE-VGGAFDLPYRPDAVLVFCGAIATL 232
Db 233 PLVVMGLSAHSDGSATVLIQQDV---GCAGLQVRKGKGGAWVPVHPVHALVNVNIGDTLEV 289
QY 233 VTGGQVKAPRHHVAAPRRDQIAGSSRTSSVFFLRPNAD 270
Db 290 LTNGRYKSVEHRAVAN-----GEQDLSLVVTFYAPAYD 322

Search completed: February 3, 2004, 17:43:36
Job time : 24 secs